Optimization of Broadband Seismic Network in the Kingdom of Saudi Arabia

Thesis by

Abdulrahman Alshuhail

In Partial Fulfillment of the Requirements

For the Degree of

Master of Science

King Abdullah University of Science and Technology

Thuwal, Kingdom of Saudi Arabia

May, 2011

The thesis of Abdulrahman Alshuhail is approved by the examination committee.

Committee Chairperson: Paul Martin Mai

Committee Member: Cvetan Sinadinovski

Committee Member: Gerard Schuster

ABSTRACT

Optimization of Broadband Seismic Network in the Kingdom of Saudi Arabia

Abdulrahman Alshuhail

Saudi Arabia covers a large portion of the Arabian plate, a region characterized by seismic activity, along complex divergent and convergent plate boundaries. In order to understand these plate boundaries it is essential to optimize the design of the broadband seismic station network to accurately locate earthquakes. In my study, I apply an optimization method to design the broadband station distribution in Saudi Arabia. This method is based on so called D-optimal planning criterion that optimizes the station distribution for locating the hypocenters of earthquakes. Two additional adjustments were implemented: to preferentially acquire direct and refracted wave, and to account for geometric spreading of seismic waves (and thus increases the signal to noise ratio). The method developed in this study for optimizing the geographical location of broadband stations uses the probability of earthquake occurrence and a 1-D velocity model of the region, and minimizes the ellipsoid volume of the earthquake location errors. The algorithm was applied to the current seismic network, operated by the Saudi Geologic Survey (SGS). Based on the results, I am able to make recommendations on, how to expand the existing network. Furthermore, I quantify the efficiency of our method by computing the standard error of epicenter and depth before and after adding the proposed stations.

ACKNOWLEDGEMENTS

A number of people have made this thesis possible. In particular I wish to thank my supervisor Dr. Paul Martin Mai who thought me the basic seismology concepts that made this work possible. I am very grateful for his tremendous support and encouragement during this project.

I would like to thank my committee members, Dr. Gerard Schuster, and Dr. Cvetan Sinadinovski from Saudi Aramco for their guidance and support throughout the course of this research.

My special thanks to Dr. Sigurjon Jonsson and Dr. Tariq Alkhalifah for their helpful discussions and valuable insights.

My appreciation also goes to my friends and colleagues and the department faculty and staff for making my time at King Abdullah University of Science and Technology a great experience. I also want to extend my gratitude to the Saudi Geologic Survey, which provided seismic data in Saudi Arabia, and to all the King Abdullah University of Science and Technology Professors and students who were willing to participate in the study.

Finally, my heartfelt gratitude is extended to my father Dr. Abdullatif Alshuhail for his encouragement and teaching me to never never never never give up.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

SGS	Saudi Geologic Survey
S/N	Signal to Noise Ratio

LIST OF FIGURES

LIST OF TABLES

Chapter 1

Introduction

Earthquakes are one of the few phenomena that can be used to understand the deep structures of the Earth. They also provide us a way to understand the mechanics of plate movements, geographic relationship between plates, and stresses acting in the crust. Earthquakes also pose a threat to humans in inhabited regions. By designing a seismic network optimally for hypocenter location one can locate these earthquakes with higher accuracy, and then be able to better study the mechanics of the plate boundaries and to define hazard maps for appropriate building codes.

The Kingdom of Saudi Arabia was considered to be seismically inactive until the 1995 earthquake in Al-Aqaba (M_L=6.4 on the Richter scale) that shook the North-West region of the country. Since then many seismic stations have been deployed, most of them located on the north-western part of the Kingdom. Three organizations were deploying the stations until 1999 when the Saudi Geologic Survey (SGS) was established and was put in charge of managing the seismic network in Saudi Arabia. Thus, the initial network was far from optimal because the three independent networks did not share their data. The seismic network is now unified under the auspices of the SGS that manages currently 81 stations with a plan to increase the number of stations to a 100 by the end of 2011.

There is only a modest amount of earthquake research being done on the Arabian plate, this can be accounted by the relative age of the seismic network and the sparse station distribution in some regions of the plate. Thus, there is need to update and introduce new stations to the network in order to make more advances in earthquake location and seismotectonics.

The objective of this work is to develop a methodology that will take into account the current network, and then suggest the optimal location for additional stations in Saudi Arabia. One can optimize for a number of different earthquake properties (location, focal mechanism, and tomography etc.), but the most essential parameter to consider is an accurate estimate of the hypocenter location. If the estimated hypocenter locations are inaccurate then all subsequent values using earthquake parameters will be degraded. Therefore, the station configuration should be primarily optimized for estimating an accurate hypocenter location. We will also condition our optimization for improved acquisition of direct and refracted waves and for increasing signal to noise to ratio of the data.

This thesis consists of the following sections. Chapter 2 presents the seismotectonic history and velocity profile of the region. Chapter 3, we describe the optimization methodology used for locating new recording stations in Saudi Arabia. Subsequently Chapter 4, illustrates the application of this methodology for updating the Saudi Arabian network and the error analysis of hypocenter location estimates before and after updating the network. Finally chapter 5 contains the conclusion.

Chapter 2

Seismological Background

The Arabian Plate shown in Figure 1 is surrounded by a number of plate boundaries exhibiting various levels of seismic activity. The seismicity, geology, and the velocity profile of the region constitute the input parameters for our procedure, to optimize the station coordinates for accurate location of regional earthquakes.

2.1 Seismotectonics of the Arabian Plate

Seafloor spreading and rifting is occurring in the Red Sea (Rodgers et al. 1999; Debayle et al. 2001; Benoit et al. 2003; Park et al 2007; Cochran 1981), which borders the western portion of the Arabian plate. The eastern plate boundary is characterized by thrust faulting due to the collision of the Arabian plate with the Eurasian plate, forming the Zagros Mountains (Stocklin 1968). This region is seismically very active and has witnessed numerous large magnitude earthquakes (Howells 1983). The depth of the earthquakes in the Zagros zone is relatively shallow, no more than 20 km in depth (e.g., Jackson and Fitch 1981; Engdahl et al. 2006). The Arabian plate boundary also comprises strike slip faults, in the Gulf of Aqabah and in the Dead Sea transform fault. The Gulf of Aqabah is associated with intense seismicity, which can be seen by the recent swarm of earthquakes that affected the North-Western Part of Saudi Arabia (e.g., Al-Amri 1995; Klinger et al. 1999).

Two distinct geologic features characterize the interior of the Arabian Plate. The Arabian Platform in the east is characterized by a sedimentary cover that thickens gradually towards the Arabian Gulf (Al-Amri, Rodgers et al. 2008). On the other hand, the Arabian Shield in the west is characterized as an igneous region associated with active volcanism (Camp and Roobol 1992). The volcanic activities (Figure 1) are accompanied by high seismic activity. Thus when looking at past earthquake maps of the Arabian Plate one can recognize an increase in seismic activity along the western border (Figure 2).

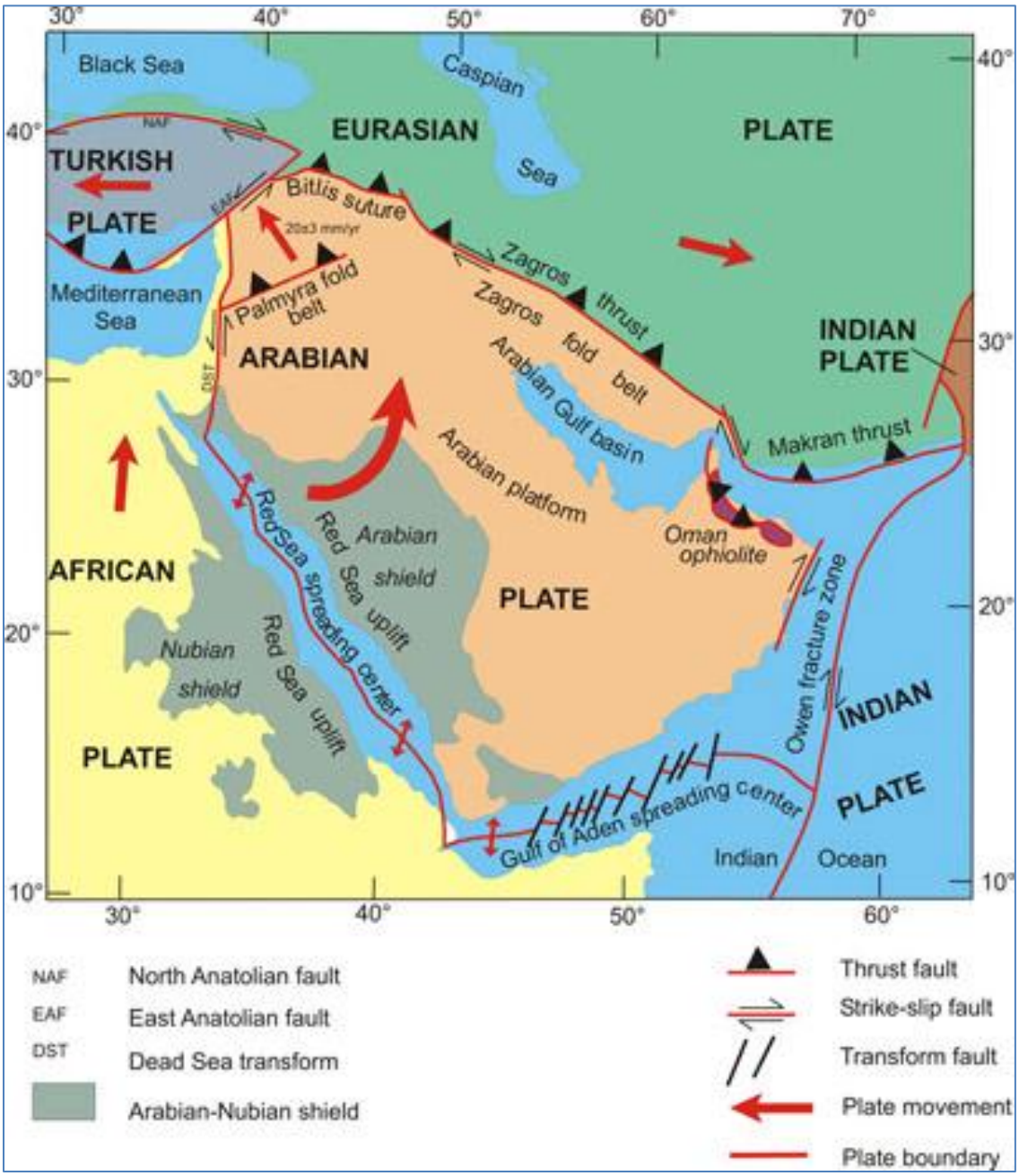

Figure 1: Schematic map of the Arabian plate boundary and the direction of movement of the plate (Courtesy of SGS).

2.2 Probability Distribution of Earthquakes in Saudi Arabia

There is no single way to define the probability distribution of earthquake locations in a region (Kijko 1977-a; Schorlemmer and Woessner 2008). For our purpose, we use seismicity maps and the regional geology to define the areas of high and low probability for earthquake occurrence. Figure 2 shows the seismic activity of the Arabian Peninsula from historic records and up to 2008 for earthquakes with magnitude 2 and larger, cataloged by the SGS. It is noticed that there are large magnitude events (M= 4-6) at the border between the Arabian plate and the Eurasian plate (in the Zagros Mountains), and a large number of events are evident at the Aqaba region although most events have low magnitudes (M < 4).

The seismicity map helps in locating the plate boundaries, and allows us to define the probability distribution of earthquake occurrence in the Kingdom of Saudi Arabia. Because there are some scattered events all over the Arabian peninsula, we added a background probability to the model. Figure 3 shows the resulting earthquake occurrence probability distribution for the same magnitude and return period of earthquakes as in Figure 2. As expected, the high values of the probability fall onto the boundaries of the Arabian plate. Regions of locally increased seismic activity in the past are mapped into higher probability than the background probability.

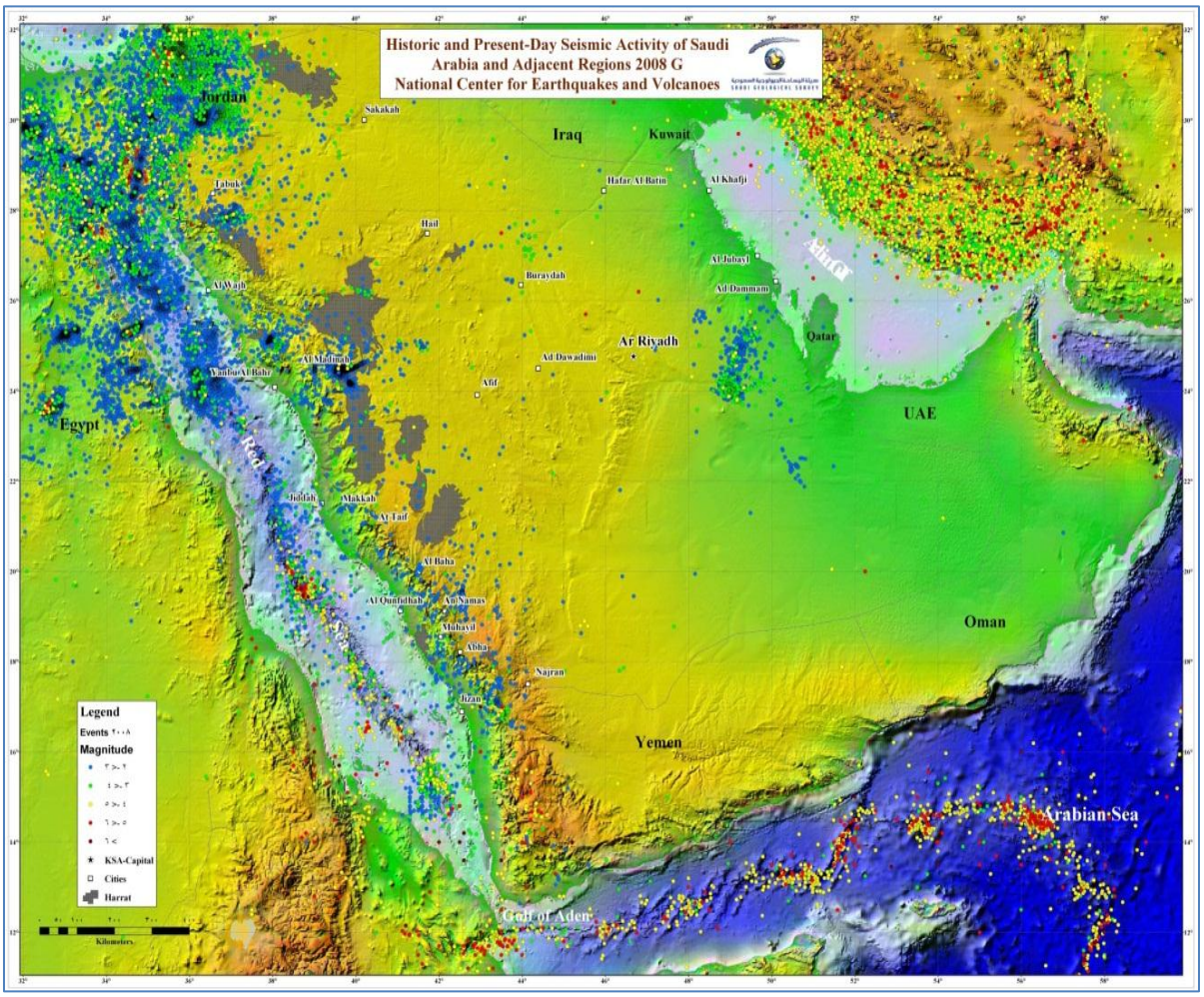

Figure 2: The seismic activity cataloged in the Arabian plate. Warmer colors indicate higher magnitude events while the cooler colors indicate low magnitude events (courtesy of SGS).

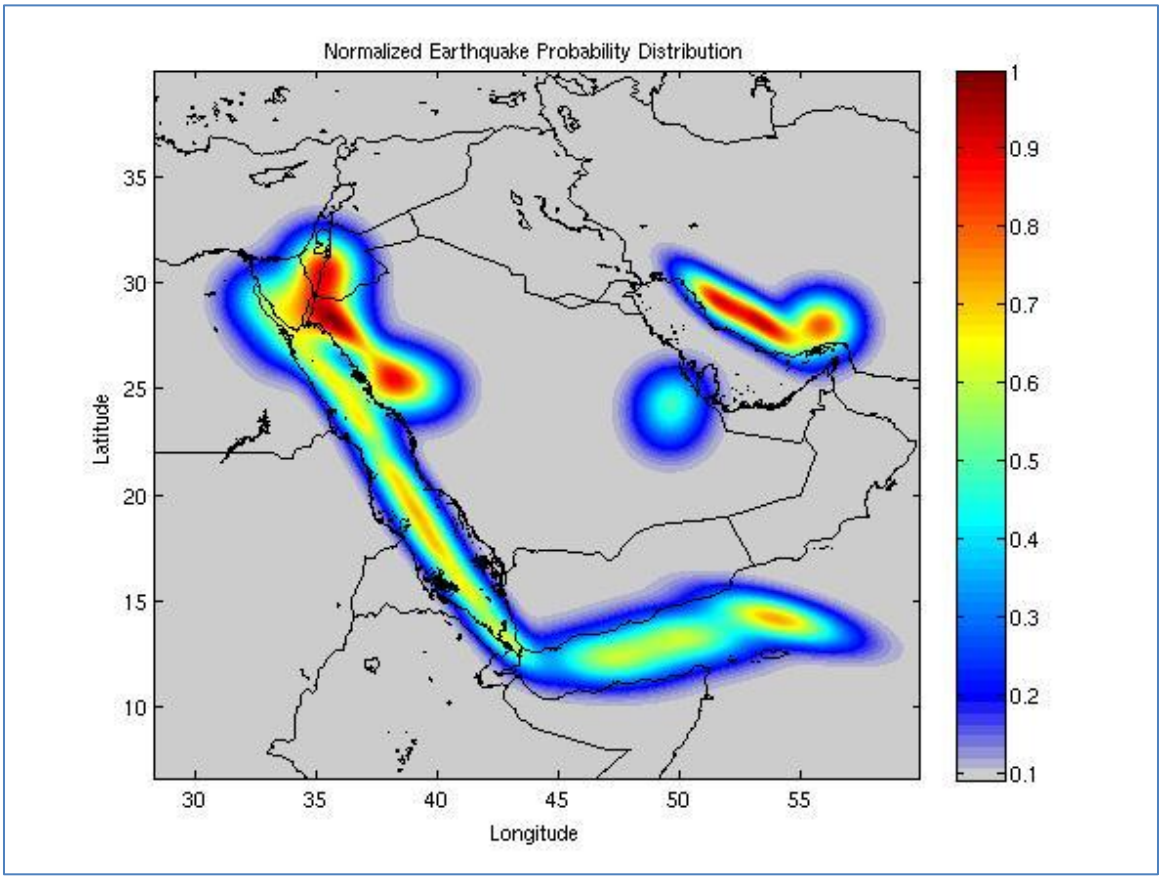

Figure 3: The normalized earthquake probability distribution for the Arabian plate. The probability is high on the borders of the Arabian plate.

2.3 Velocity profile of Saudi Arabia

For our optimization approach, an accurate P-wave velocity model of the Arabian Peninsula is needed. Waveform modeling was used in previous studies (Rodgers, et al. 1999) to model the 1-D velocity profile of both the Arabian shield and Arabian platform (Table 1). They used data obtained from the 1995-1997 Saudi Arabian Temporary Broadband Deployment (9 stations), and calculated synthetic seismograms to fit the observed data using a grid search scheme. Another 1-D velocity profile for the Arabian

Platform was obtained by Al-Amri (1998) using spectral analysis of long period P-wave amplitude ratios (Table 2). He selected 11 earthquakes to calculate the theoretical P-wave motion and compare it to the observed spectral ratios.

Table 1: Velocity Profile Obtained from Waveform Modeling.

Arabian Shield Thickness (km)	P-wave Velocity (km/s)		Arabian Platform Thickness (km)	P-wave Velocity (km/s)
1	4.0		4	4.0
15	6.2		16	6.2
20	6.8		20	6.4
Half-space	7.9		Half-space	8.1

Table 2: Velocity Profile Obtained by Spectral Analysis.

Layer	P-wave Velocity (km/s)	Depth (km)
1	5.6	3.0
2	6.3	13.0
3	6.6	21.0
4	6.9	36.0
5	7.6	46.0
6	8.3	Half-space

Because we first want to use a simple and unifying approach, we deploy a single velocity model. Additionally, we need to calibrate our method against the current station deployment of the SGS. Thus, we adopt the velocity model shown in Table 3.

Table 3: Velocity Model for Saudi Arabia

Layer	P-wave Velocity (km/s)	Depth (km)
1	4	0-4
2	6.2	4-20
3	6.4	20-40
4	8.1	>40

Alternative models can be tested, and as a new 1D-referance model becomes available or a locally different 1D model is available for a localized study, the method can be adapted. In principle a 3D velocity model should be used for hypocenter location, as more velocity information is known and made available one can refine the method for a 3D velocity model.

Chapter 3

Optimization Theory

A number of different methods have been developed for network optimization. Kijko optimized the network for hypocenter location (Kijko 1977-a). Another method was developed to design the optimum network for aftershock recordings (Hardt and Scherbaum 1994), it uses the same basic theory as Kijko's method but applies a number of different constraints to better suit the problem of designing aftershock networks. Another method has also been optimized for early earthquake warning system in Turkey (Oth et al. 2010). Because each has different criterions for optimization, one should choose a suitable method depending on the objective of the network. In our case, our primary objective was to optimize the network for hypocenter location.

3.1 Strategies for Network Optimization

Our method for optimization of seismic station distribution was originally developed by Kijko (1977-a), and provides a comprehensive and effective means for optimizing spatial location of recording stations for accurate location of regional earthquakes (Ghalib, Russell et al. 1984). It is based on "D-optimal planning criterion" (Box and Lucas 1959). That seeks the best station locations that minimize the covariance's of the unknown earthquake locations in a least square sense. For hypocenter location, observed arrivals

times recorder at a set of seismic stations, are compared to corresponding predicted values.

The standard least squares problem for earthquake location is defined as finding the hypocentral location $\theta = (x_o, y_o, z_o, t_o)$and initiation time that minimize the sum of the squared data residuals:

$$min_\theta = \sum_{i=1}^{N} (t_i^{obs} - t_i^{pred}(\theta))^2, \tag{3.1}$$

where t_i^{obs} denotes the observed arrival traveltime of a P-wave at station *i*, t_i^{pred} denotes the calculated arrival time at station *i* as a function of θ for and assumed velocity model.

For an assumed distribution of earthquake locations we are seeking the optimal station network that will minimize the error in calculating the hypocenter parameters. For a set of seismic stations given by $X = (x_i, y_i, z_i)$ the computed arrival time to the *i*th station will be a function of station location, X_i, and hypocenter location, θ_i:

$$t_i^{pred} = f(X_i, \theta). \tag{3.2}$$

$\delta\tau$ is the weighted difference between the observed and computed arrival time, also known as the data residual, and is given by

$$\delta\tau_i = t_i^{obs} - f(X_i, \theta) \tag{3.3}$$

Thus, the matrix of partial derivatives for the computed arrival times with respect to the hypocenter location will take the form

$$A_{ij} = \frac{\partial f(X_i,\theta)}{\partial \theta_j}. \tag{3.4}$$

This is equivalent to linearizing the nonlinear problem. For four or more stations the over determinate weighted least squares solution to the hypocenter location is expressed as (Menke 1984):

$$\delta\theta = [A^T W^{-1} A]^{-1} A^T W^{-1} \delta\tau. \tag{3.5}$$

$\delta\theta$ is the perturbation in hypocenter parameters and W is a diagonal matrix of arrival-time variances, in principle W should be assumed or calibrated from observations.
In this expression, the covariance matrix of hypocenter parameters (Menke 1984) is defined as

$$D(X) = [A^T W^{-1} A]^{-1}. \tag{3.6}$$

Assuming a constant data variance for all stations (Draper and Smith 1966), then:

$$W = \sigma^2 I \tag{3.7}$$

the resulting model covariance matrix will be:

$$D(X) = \sigma^2[A^T A]^{-1} \tag{3.8}$$

The goal is to select a station distribution (X) such that the determinant of the model covariance matrix D(X) will be a minimal (Box and Lucas 1959). This condition will minimize the confidence volume for the hypocenter parameters. It can be stated as "X' is D-optimal if and only if D(X') is nonsingular and min DET[D(X)]=DET[D(X')]" (Ghalib, Russell et al. 1984), where DET stands for taking the determinant. Using the properties of determinates one does not necessarily have to compute the covariance matrix (Ghalib, Russell et al. 1984), which reduces the computational cost. Therefore, the trace of A^TA is the sum of the positive eigenvalues. One desires a well conditioned inverse matrix $(A^TA)^{-1}$,hence a well conditioned A^TA, thus requiring that the sum of the positive eigenvalues is small. The inverse of A^TA is inversely proportional to determinant of A^TA. Hence, one desires a large determinant to keep values of $(A^TA)^{-1}$ small.

The operator (L) can be defined such that

$$L[D(X)] = DET[D^{-1}(X)] = DET[A^T A], \tag{3.9}$$

Equation 3.9 assumes a constant variance of unity (σ^2=1), but the constant values assumed for the variances could be relaxed. The data variance is a statistical property that depends on several factors (background noise level and accuracy of seismic station instruments), for optimizing the geometrical relations between station distribution and

hypocenter location the constant was set to unity. Therefore, by maximizing L[D(X)] the determinant of the model covariance matrix will be minimized.

In the case presented above the optimization of stations distribution was carried out for a single event (one set of hypocenter parameters). Kijko (1977-a) expanded on this and included stations distribution for a set of variable hypocenters $\theta' = (x_O, y_O, z_O, t_O)$ with a normalized probability distribution given by $p(\theta')$. The ensemble averaged L-operator integrated over the values of the random variable is given by

$$\langle L[D(X)]\rangle = \int_{\theta'} p(\theta')L[D(X,\theta')]d\theta'. \tag{3.10}$$

The above expression is equivalent to averaging the values of the determinant over all possible events in a region. Thus, maximizing $\langle L[D(X)]\rangle$ (our chosen cost function) will give the optimal station distribution in an average sense. For a discrete distribution of hypocenters the ensemble average can be approximated as

$$\langle L[D(X)]\rangle \approx \sum_{k=1}^{n} p(\theta'_k)L[D(X,\theta'_k)]\Delta\theta'_k\ , \tag{3.11}$$

where n is the number of stations used to optimize the network. The accuracy of event location depends on an accurate velocity model (to compute predicted arrival times), and a precise earthquake probability distribution. Unfortunately, both these quantities are very difficult to acquire and approximations must be made for the region of interest.

3.2 Forward Modeling

The optimization method was tested first on a simple case, of constant velocity and a single hypocenter. We evaluated the forward model for a single hypocenter with location (x,y,z)= (12.2 km, 12.5 km, 8 km) and optimizing the station distribution for four additional stations. This was done by setting trial station points every 1 km in each axis. The results (Figure 4) display three stations covering the epicenter from equal azimuthal directions in order to resolve the X and Y components, while one station is located close to the epicenter to constrain the depth of the event.

This arrangement corroborates the physical intuition as the stations cover the epicenter equally from all possible azimuths using three stations while reserving one station to constrain the depth such that they optimally resolve the x, y, and z components. This also agrees with the triangular quadripartite network recommended by Lilwall and Francis (Lilwall and Douglas 1970) and Uhrhammer (Uhrhamer 1980).

This configuration is perhaps useful for networks where earthquakes are likely to occur within the network. Most seismic events originate on the borders of the Arabian plate. Coverage of the seismic events from all directions will be restricted, due to the borders of Saudi Arabia being limited inside the Arabian plate. Thus, one needs to test for configurations where the event is located outside and on the border of the network. Figure 5 shows another hypothetical case that was generated where the hypocenter location is (x,y,z)= (12.5 km, 25 km, 8 km) and four stations were optimized. This simulation represents an event happening on the border of the potential network.

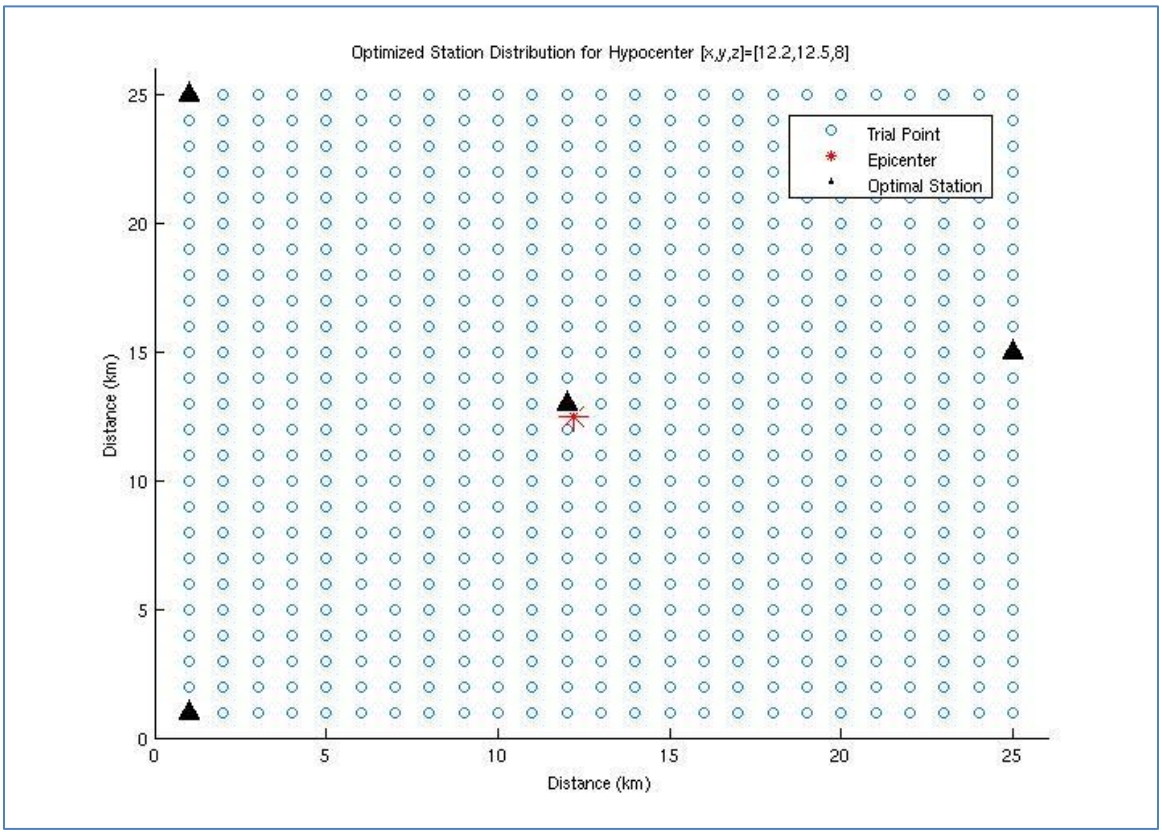

Figure 4: Forward model distribution case-A. The red star represents the earthquake epicenter while the black triangles represent the optimal station configuration, the blue circles represent the trial points.

Analyzing Figure 5 we find that the optimum configuration tries to cover the event form all different azimuths within 180 degrees of degrees of freedom, while keeping one station near the event in order to estimate the depth accurately.

Two more hypothetical situations were tested where the seismic event originated outside the network. Similar results were obtained but the azimuthal restriction in covering the event is more severe.

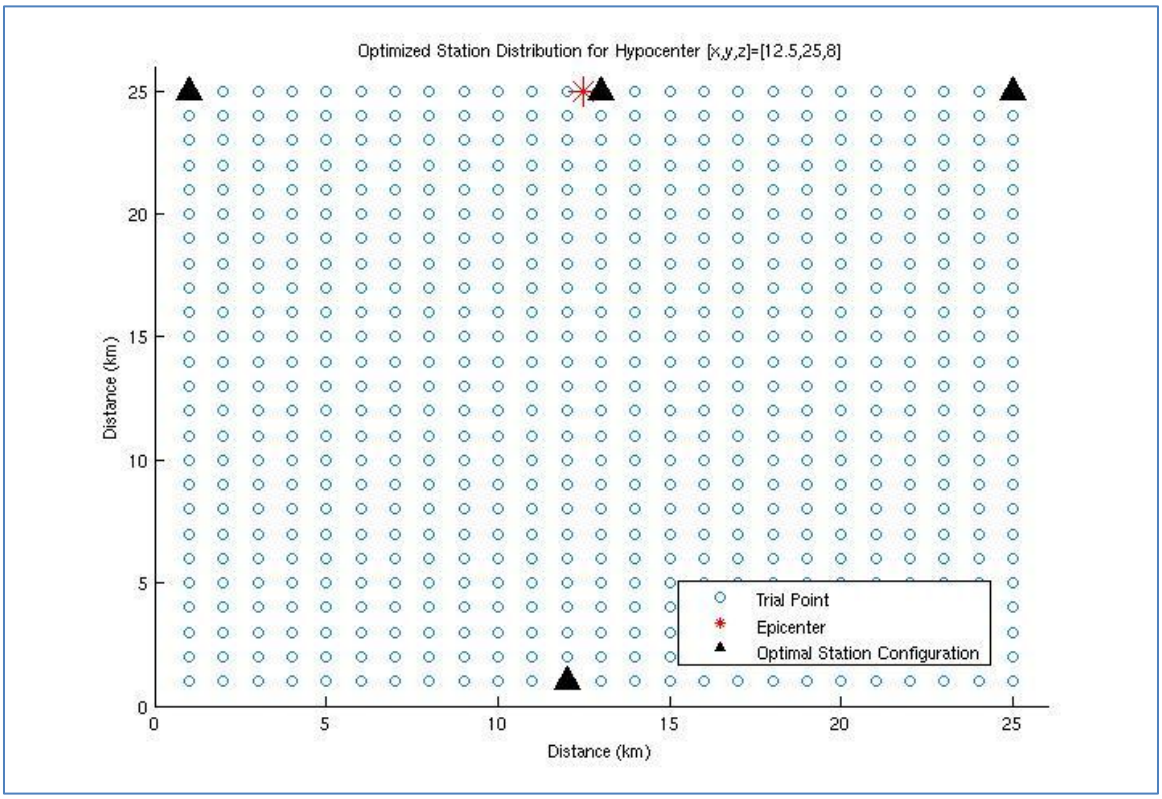

Figure 5: Forward model distribution case-B. The red star represents an earthquake epicenter while the black triangles represent the optimal station configuration; the blue circles represent the trial points. Since the event is on the border the freedom of covering the event is severely restricted.

To understand the effect of depth of the seismic event on the near station distance to the epicenter, I varied the depth from 3 km to 13 km (Figures 6 and 7). Notice how the near station tends to move farther away from the seismic event for the deeper earthquake, because the closest station will now be able to resolve the depth even at farther distances. In this case, the near station was integrated to some extent into the far stations, hence coverage is now achieved from four azimuths rather than three in the previous cases.

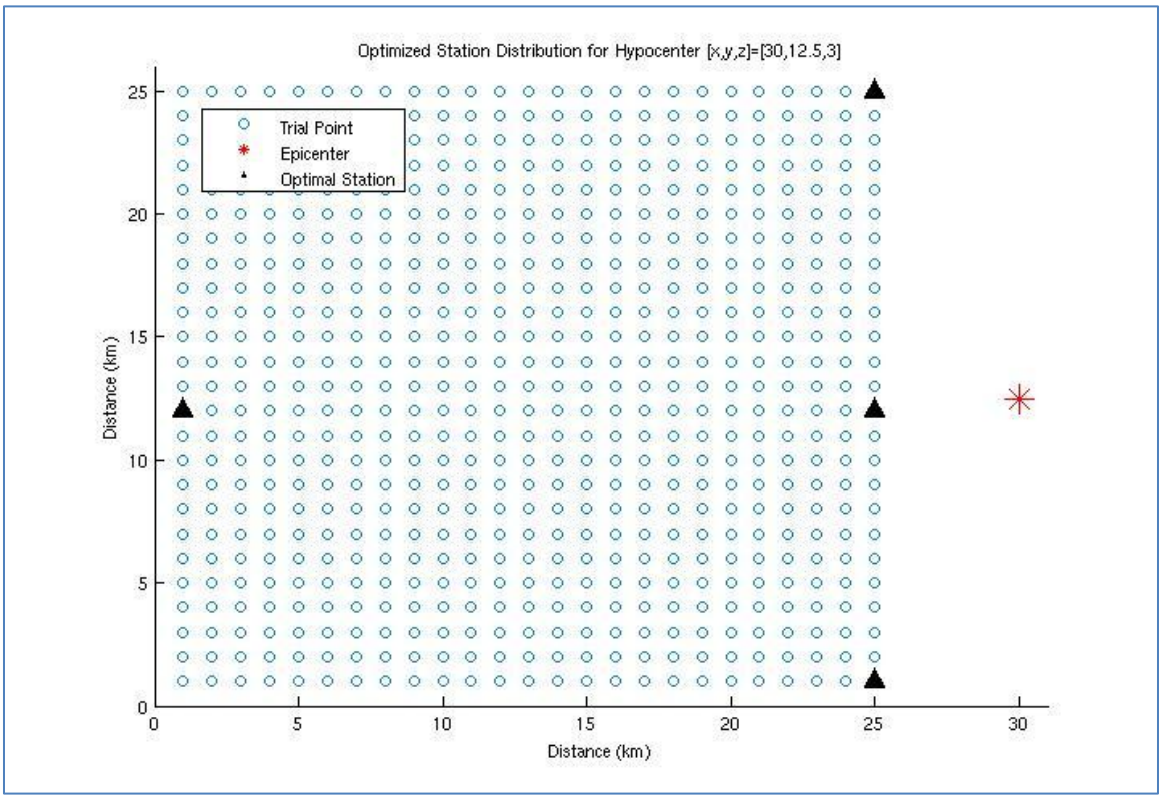

Figure 6: Forward model distribution case-C. The epicenter is located outside the boundaries of the network with a shallow depth of 3 km. Azimuthal coverage is achieved only inside the limited network.

An interesting feature of the cost function is that it tends to place the stations as far as possible from the seismic event due to the geometry of the solution. This is explained by Rabinowitz and Steinberg (1990), stating that the D-optimal criterion is a monotone increasing function of the sine of the take-off angle of the ray path from the hypocenter to the station. Thus, the maximum will occur at the farthest distance in the grid from the epicenter (Rabinowitz and Steinberg 1990).

The arrangement above does not consider geometric spreading, which in fact will yield data recorded with low signal to noise ratios, therefore creating an uncertainty in hypocenter estimation.

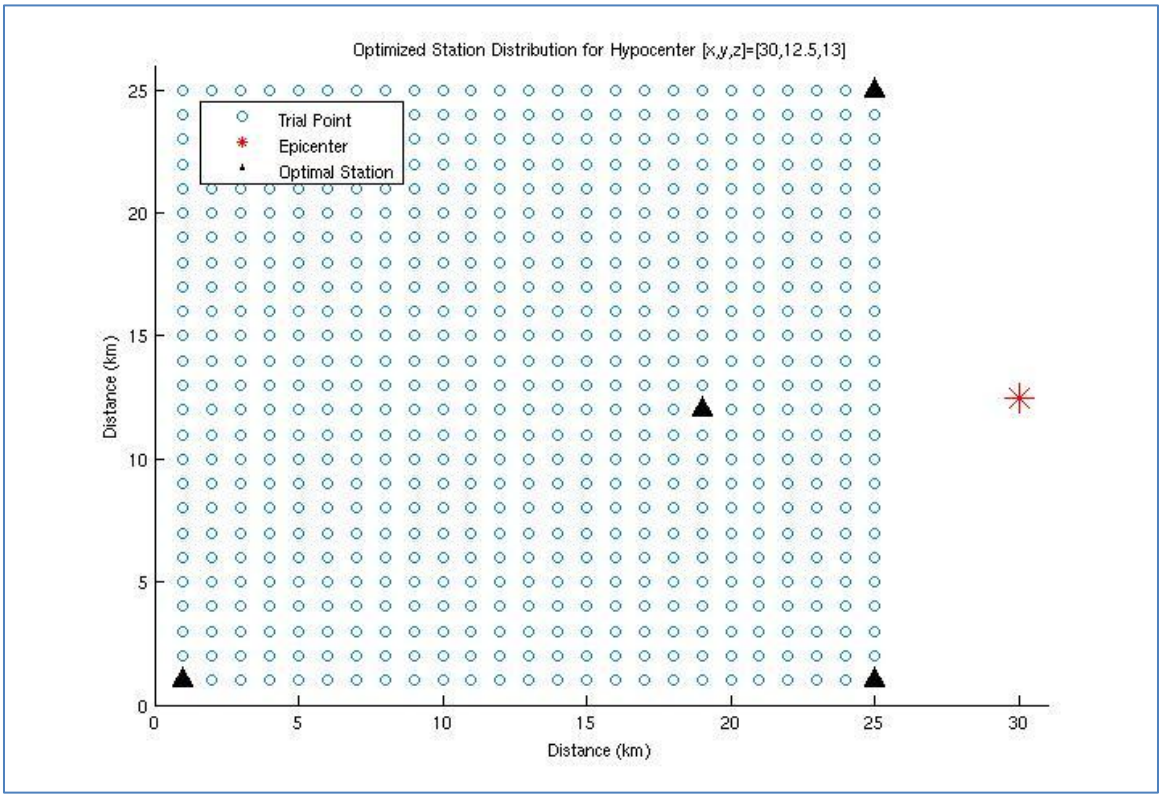

Figure 7: Forward model distribution case-D. The epicenter is kept the same as in Figure 6 but the depth is drastically changed to a deeper event (13 km). Notice how the near station tends to move farther away from the epicenter.

3.3 Accounting for Geometric Spreading

The previous examples show that the stations tend to move to the edges of the network. To reduce this effect, we modify the cost function, by including a geometric spreading term such that the cost function value decreases with distance.

For a homogenous elastic solid, body wave amplitudes decease as 1/r due to geometric spreading (Stein and Wysession 2003). For a layered medium with an increase in velocity with depth, refracted wave energy reaching the surface shows a sudden increase in energy just beyond the crossover distance(defined as the distance at which rays penetrating into the deeper layers and having longer ray path arrive simultaneously with the rays propagating only in the top-most layer). A layered medium with a lower velocity zone generates a distance interval in which no rays arrive at the surface (so called shadow zone).

To account for the geometric spreading effect the cost function was designed such that it decreases as 1/r for a homogenous medium because the amplitude also decreases as 1/r. This places the stations at reasonable distances from the hypocenter. In case of a model with increasing velocity, one can increase the value of the cost function for an interval just beyond the crossover distance. Thus, the stations will record the refracted arrivals. In case of a decrease in velocity within a layer, the cost function can be reduced dramatically in value at the shadow zone. This decrease is kept constant until the distance in which rays start arriving again. Figure 8 displays typical ray paths in layered media and the corresponding cost functions (Stein and Wysession 2003).

The effect of this cost function penalty is to place the stations to be closer to the epicenter. If velocity increases with depth, stations are located such that they optimally record both direct and refracted waves.

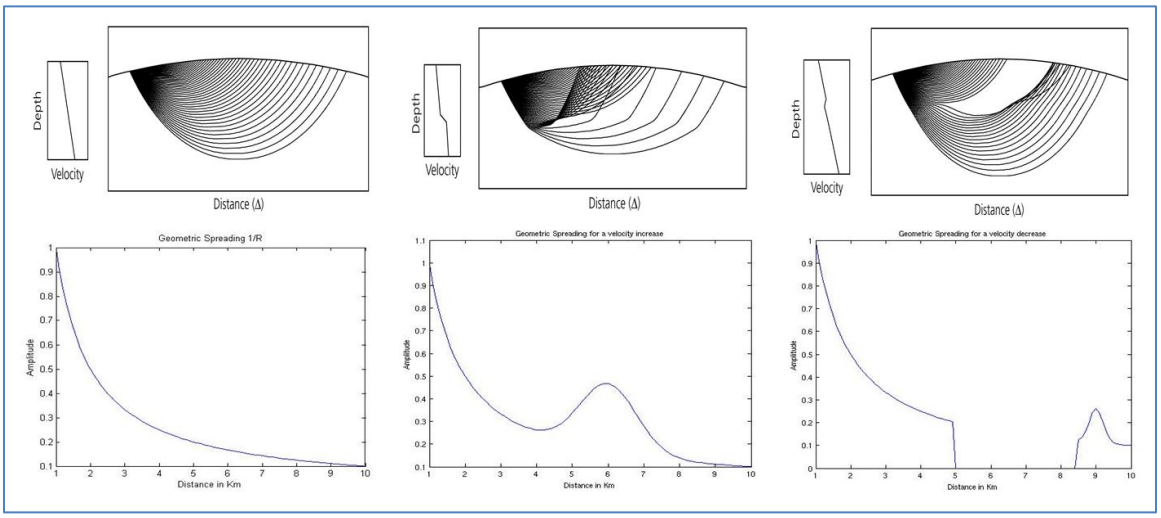

Figure 8: Cost function penalty for geometric spreading. Top panels display ray diagrams for velocity models (Stein Wysession 2003). Bottom panels show the penalty factor that is to be applied to the cost function.

3.4 Station Clustering

For multiple tests with a single hypocenter and more than seven stations to optimize, we notice that two stations are located right next to each other. This effect is documented in previous papers (Rabinowitz and Steinberg 1990), and is related to the assumption that model errors are statistically independent. It also arises from the design of the problem, if many stations are being optimized simultaneously for a single event. Since station clustering will not improve the performance of the network, we attempted to overcome this effect.

There are three solutions to this problem. The first is not to assume that the errors from the model are statistically independent, therefore σ^2 is not the true identity as defined in Equation 3.7. Another solution is to optimize the network for multiple events simultaneously, i.e. to introduce a probability distribution of earthquakes. This tends to

distribute the stations more widely. Finally, one can introduce a penalty for stations that are placed within a certain distance of each other. While applying the method we used 80 km as the minimum distance between two stations.

Our objective is to improve the network design for earthquake location in general, hence we optimized for a probability distribution rather than one source, and included a penalty for stations within 80 km next to each other.

3.5 Optimization technique (Pattern Search)

Pattern search is considered to be a direct search method (Torczon 1997). It is a derivative free method that will find the maximum of a function by analyzing how the function behaves with different directions in a “multidirectional search algorithm”. It starts out with an initial value for the parameters to be estimated, then searches in multiple directions until it finds where the cost function is decreasing, and then iterates based on decreasing the cost function even lower (Torczon 1997).

Pattern search does have its limitations, it only works for well behaved cost functions that are not highly non-linear. To test the general behavior of our cost function, we kept all the stations fixed, except for one, and evaluate the cost function at each trial point.

Figure 9 shows the cost function behavior for four stations. The three outer stations are kept fixed while the middle station is moved. The arrangement of the outer stations is exactly the same as in Figure 4. The cost function shows a maximum at the location of

the epicenter, a physically plausible solution because this is where the fourth station needs to be placed in order to constrain the depth. Figure 10 shows the cost function behavior for four stations, where only one of the outer stations was allowed to change its locations.

The solution (hypocenter location) for an event will be unique if there are exactly four stations recording its arrival times, assuming no errors and sufficient S/N ratio. This is because the error ellipsoid will intersect at one point. Adding additional stations will not improve the accuracy of the solution (unless there are errors in the arrival times), and add to the non-uniqueness of the solution as it will become an over-determined inverse problem. If there are errors in the measurements, the solution will be an area or volume, rather than a point and by adding stations its area or volume will be decreased. To visualize this we calculate the mean error associated with the hypocenter parameters xo,yo,zo (see section 4.2 for the theory) for different number of stations to optimize for (Figure 11). We notice that the error decreases rapidly at first, and then seems to stabilize, suggesting that the increasing the number of stations will not affect accuracy significantly past a certain stage.

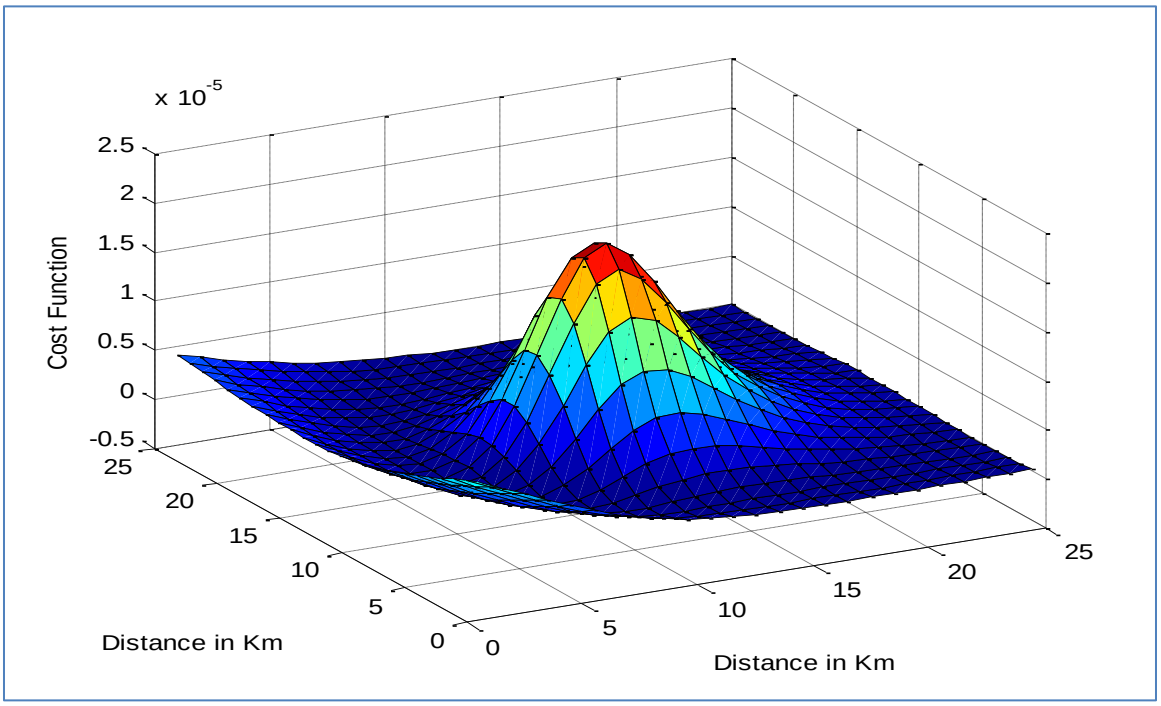

Figure 9: Cost function behavior case-A. The cost function, with three outer stations kept fixed. Notice the maximum coincides at the epicenter (12.5 km, 12.5 km) in order to constrain the depth of the event.

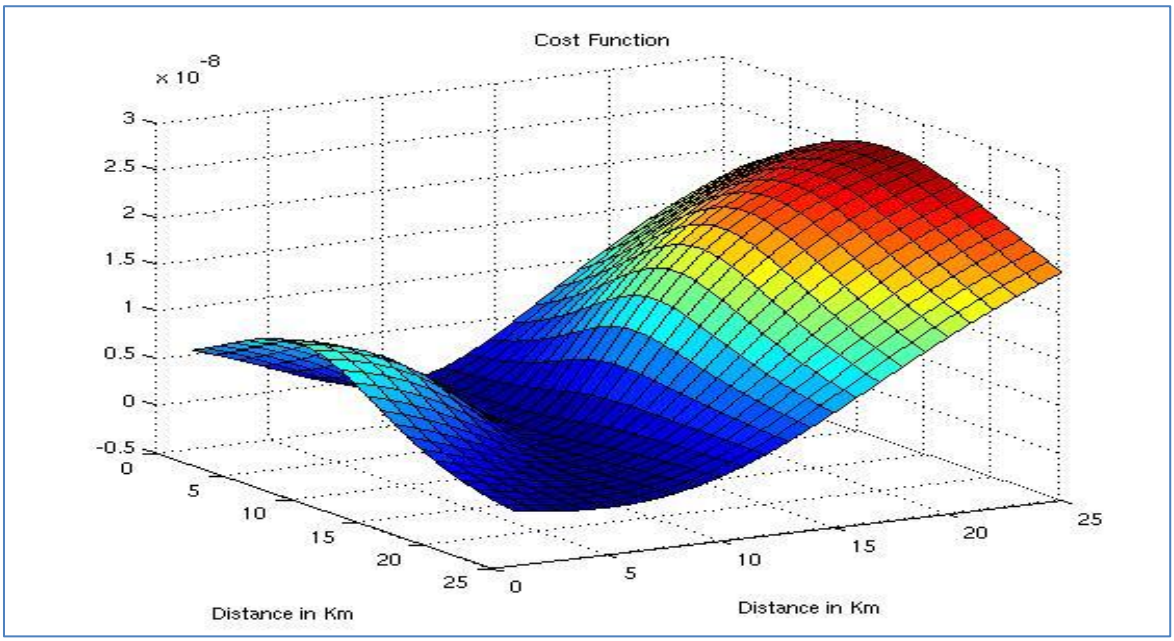

Figure 10: Cost function behavior case-B. The cost function, with two of the outer stations kept fixed while the third was not kept fixed. Notice the maximum coincides at the edge to constrain the x and y components.

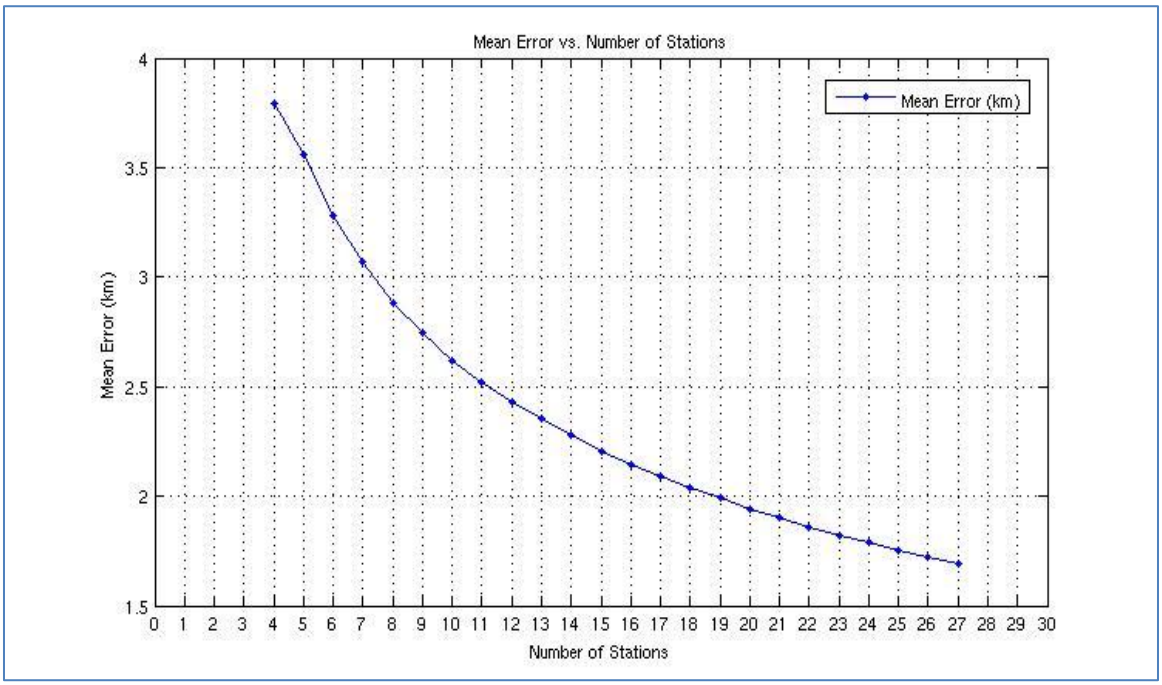

Figure 11: The smoothed mean error of hypocenter parameters (xo,yo,zo) in km versus the number of stations used for optimization of locations.

3.6 Reliability test

To perform a reliability test of the pattern search algorithm, we apply the optimization procedure to the hypothetical case of a four station network to locate a single event located at (x,y,z)=(50 km, 50 km, 10 km) (Figure 12). The velocity model used is the one in Table 3. The optimum network places three equidistantly spaced stations in a large outer circle around the epicenter, while one station is placed very near to the epicenter. This corresponds to the theoretical optimum network the “triangular quadripartite network” (Uhrhammer 1982; Rabinowitz and Steinberg 1990). Figure 12 shows the results for the hypothetical case, where the red star represents the epicenter of an

earthquake with a depth of 10 km while the black triangles represent the optimum station locations.

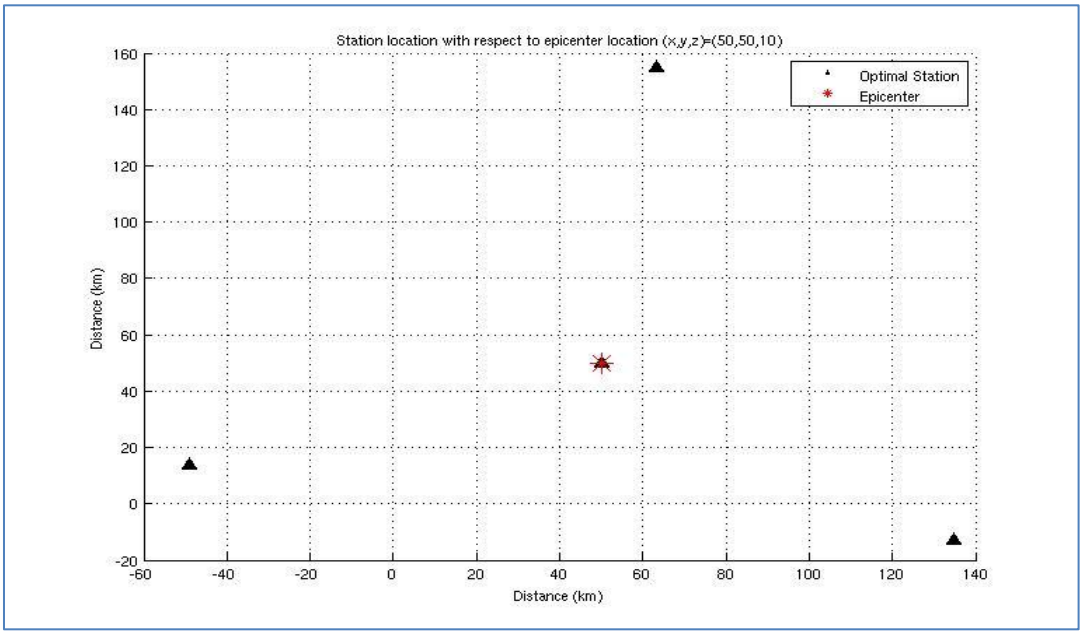

Figure 12: Optimization of four stations for a single source. The red star represents the epicenter and the triangles represent the optimum station configuration.

We repeat the previous test, optimizing for seven stations to locate an earthquake that originates at (x,y,z)=(50km, 50km, 10km).

The optimized network places the stations at two concentric circles (that have different radii) centered at the epicenter. The stations at the smaller circle record the direct wave (due to their offset being less than the crossover distance) and constrain the depth. The stations in the outer circle are located beyond the crossover distance where refracted rays start arriving to recover the high amplitude refracted wave energy beyond the crossover distance. This enables the network to accurately locate the

hypocenter, but also facilitates refraction tomography studies to image crustal structures. Figure 13 shows the results for this hypothetical case, where the red star indicates the epicenter while the black triangles indicate the optimum station configuration.

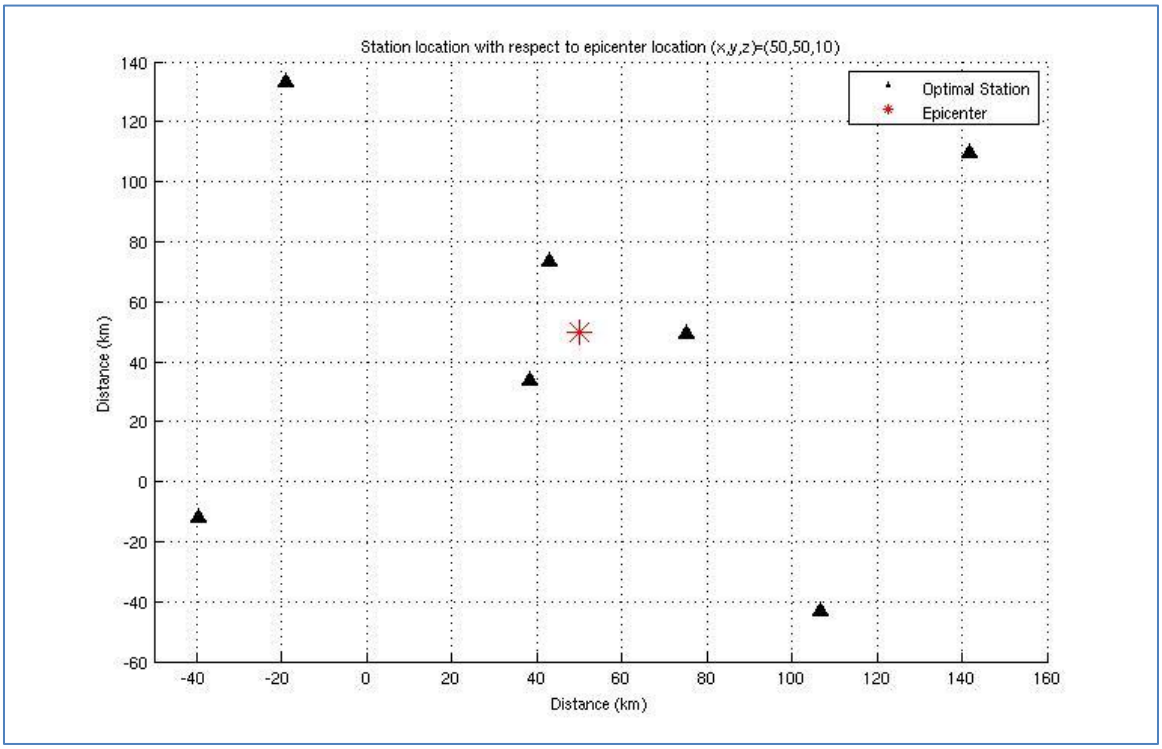

Figure 13: Optimization of seven stations for a single source. The red star represents the epicenter and the black triangles represented the optimum station configuration.

Analyzing the convergence of the solution (for the case presented in Figure 13), we notice that the solution converges in less than two hundred iterations. Figure 14 displays the cost function change with iteration number. The convergence rate depends on two factors: the initial station distribution and the number of stations to optimize.

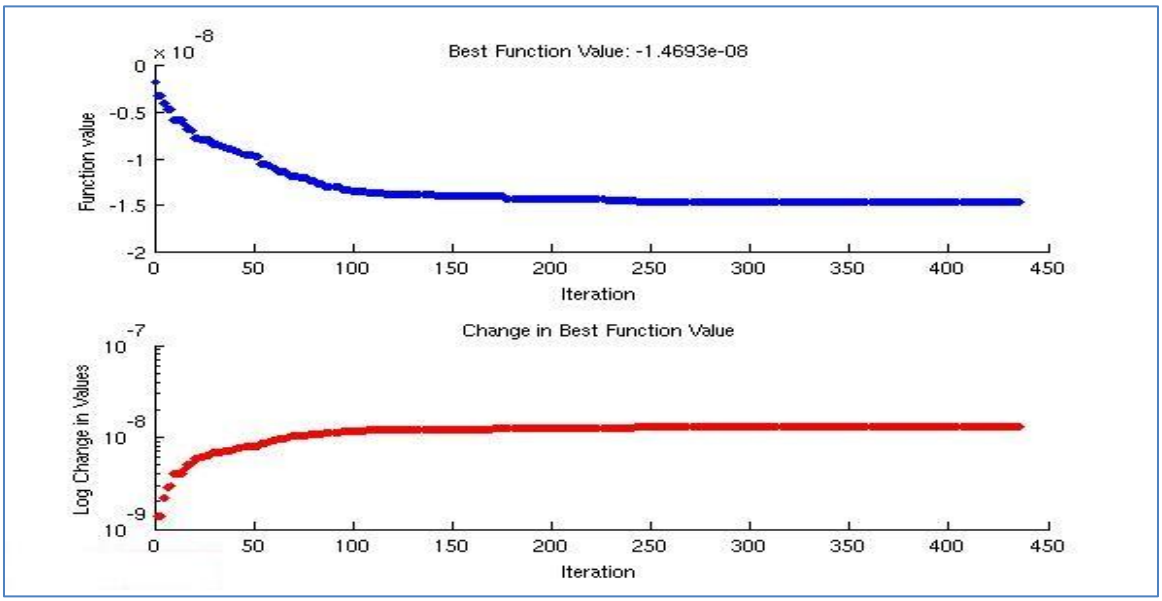

Figure 14: Convergence of the cost function. Top: cost function value versus iteration. Bottom: Log change in cost function values versus iteration.

3.7 Probability Distribution of Earthquakes

In all previous cases we considered a single earthquake for designing the network. However, in nature an earthquake can originate within a large area with different probability of occurrence in space and time. By considering a probability distribution, we effectively average the value of the cost function over all possible events in the region (Garciafernandez, Kijko et al. 1988). This was tested on a simple probability model of a small square with high earthquake probability, surrounded by a larger square with a low earthquake probability. Figure 15 shows the results for the optimized station locations. This test illustrates how the method arranges the stations along two circles with same center and different radii with close to equal azimuths.

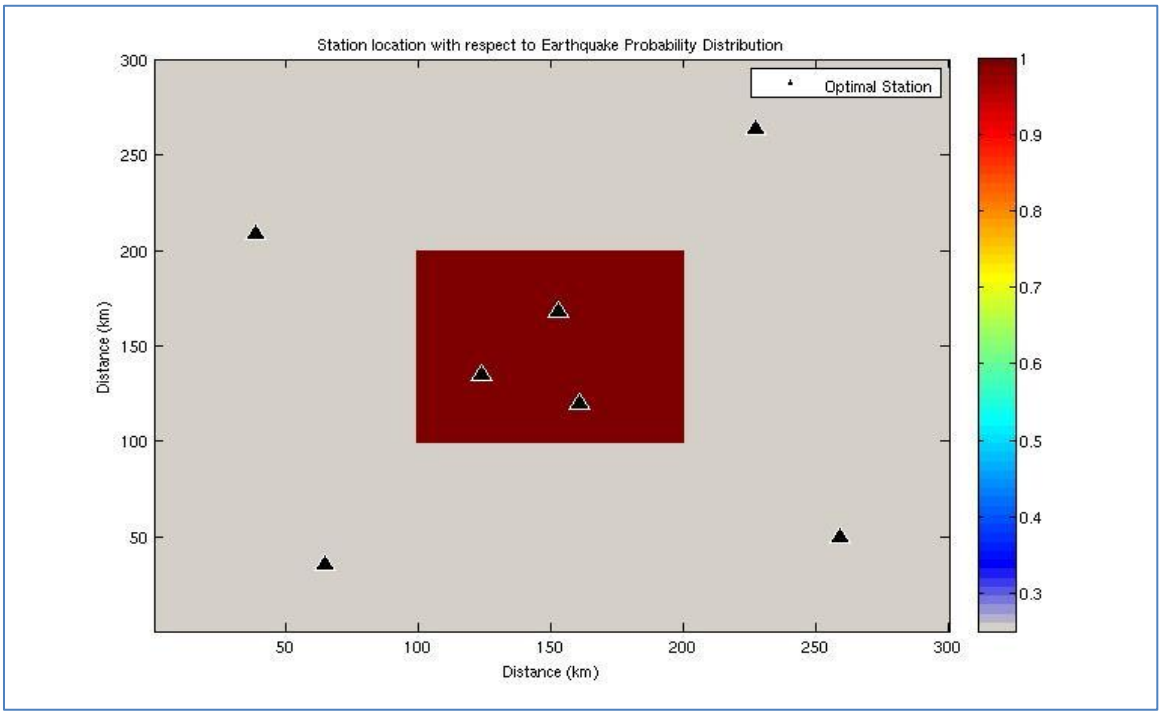

Figure 15: Optimization of seven stations for a probability distribution (Square). The black triangles represent the optimal station locations for the given earthquake probability model.

While such simple shapes certainly do not resemble the underlying geology, the chosen probability of earthquake occurrence should also change gradually rather than abruptly (like in the previous example). Since earthquakes usually happen along faults, the probability should have high probability values along the fault, and then slowly decrease orthogonal to the fault. We use a Gaussian distribution to approximate such a probability distribution along a fault (Figure 16). Using this probability distribution in our optimization code, we obtained the station distribution in Figure 17; stations are aligned in two concentric circles with different radii, while keeping the azimuth spacing almost equal.

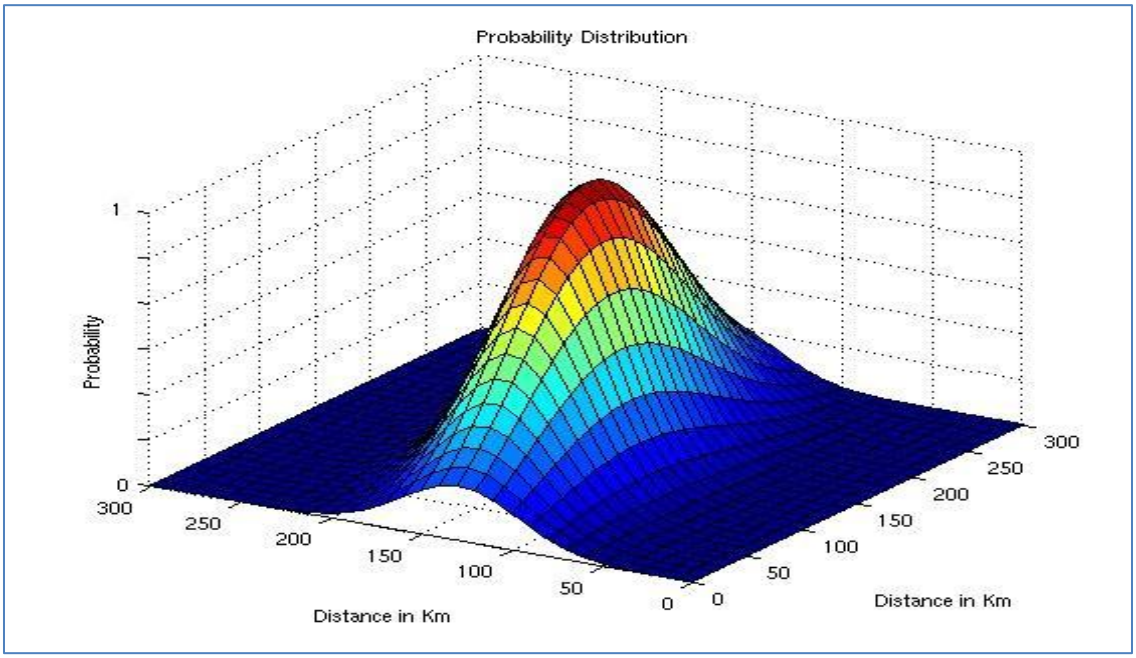

Figure 16: Hypothesized earthquake probability distribution along a fault.

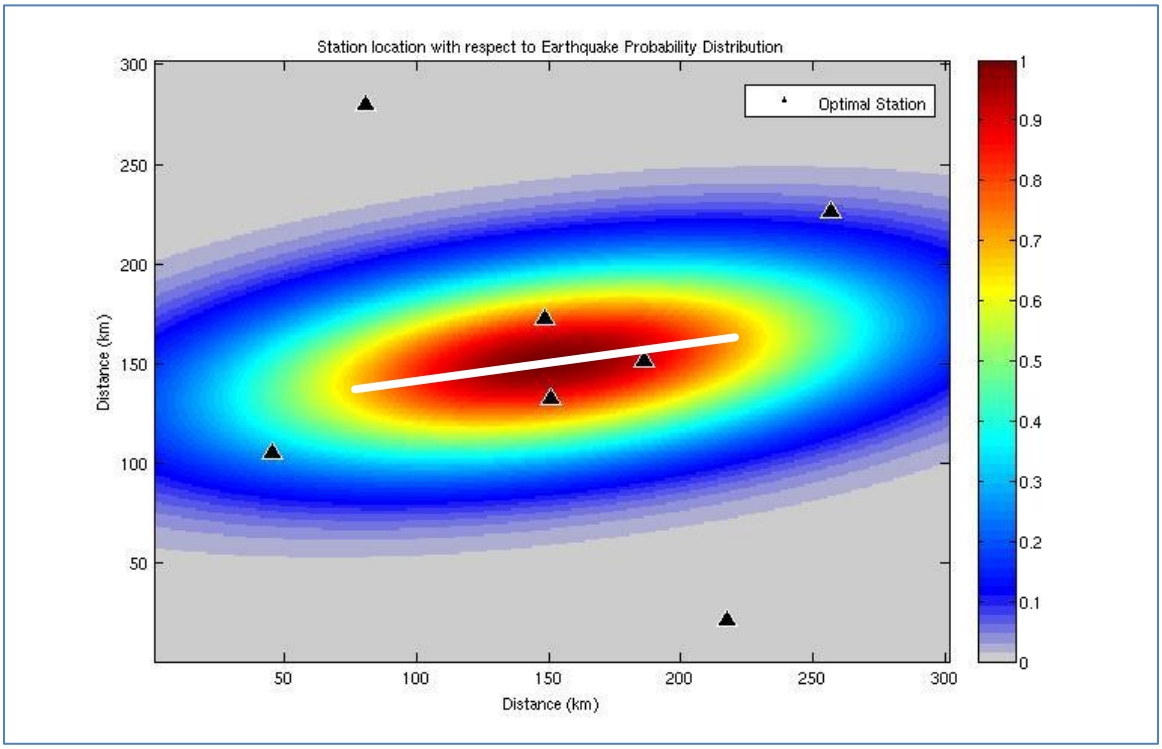

Figure 17: Optimization of seven stations for an earthquake probability distribution along a fault. The optimized stations (black triangles) are placed around the fault (white line).

In most of the cases an existing network needs to be updated by adding stations. Thus, we also test our method for updating an existing network. Figure 18 shows an assumed existing network (Brown circles) of six stations updated by five optimized station locations (Black triangles). The optimization places the new stations such that the average creates the two concentric circles with different radii.

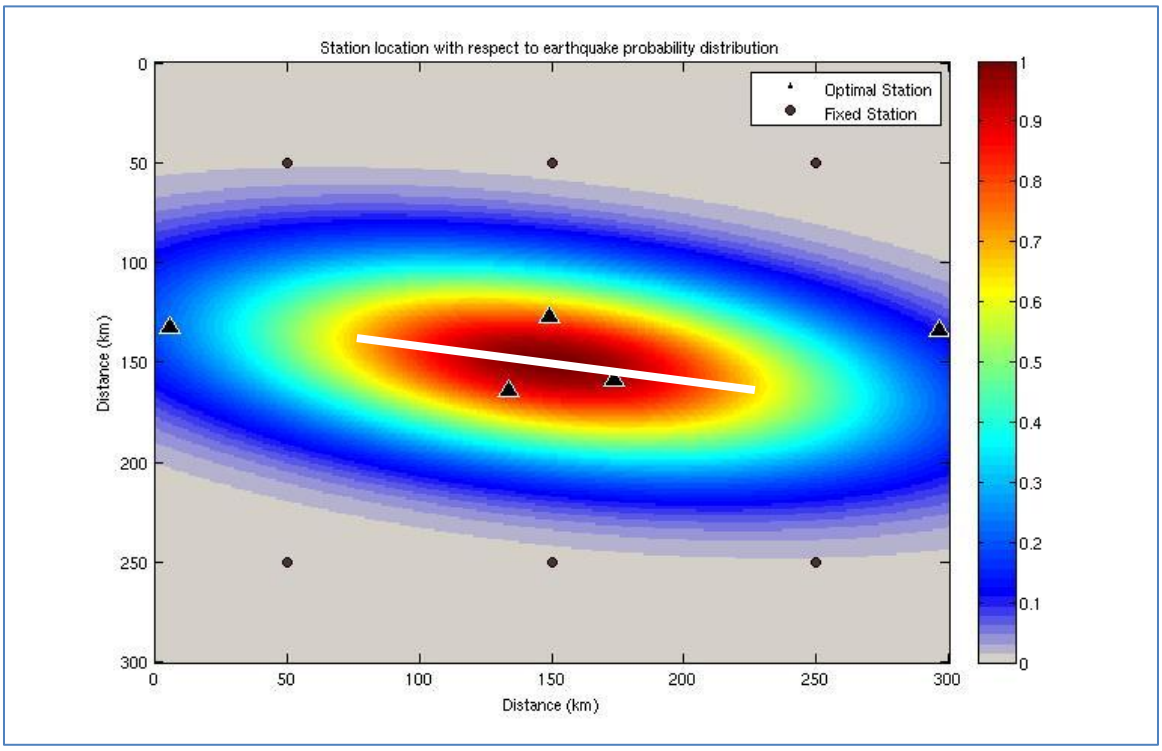

Figure 18: Optimization of five stations for an existing network.

Chapter 4

Application to the Kingdom of Saudi Arabia

In the following, we apply the methodology presented in this thesis to update the Saudi seismic network by expanding the station distribution and adding more stations to obtain more accurate hypocenter locations. Improved maps of the seismic activity in turn help in understanding the seismotectonics of Saudi Arabia.

4.1 The Optimized Saudi Network

Using the normalized earthquake probability distribution (Figure 3), the selected velocity profile (Table 3), and the existing SGS stations distribution we apply the technique to define the optimum station locations. In this particular case, the constraint applied is that all the stations had to be inside the borders of Saudi Arabia. Although most of the large seismic events happen outside the Saudi boundary, this makes the coverage from the Saudi side very limited. We configured the optimization to update the current SGS network with six stations (Figure 19). Table 4 shows the longitude and latitude values for the optimal station locations. The optimized stations are placed along a curved line that runs from north to south in central Saudi Arabia, This line curves towards high earthquake probability zones. In the north the line is concave towards the Aqaba region whereas in the south it is concave towards the eastern region, where the Arabian plate meets the Eurasian plate. Because the zones of high earthquake

probability are outside Saudi Arabia and on opposite sides (Zagros faults in the east; Red sea in the west) the optimization placed the stations in the center of the country to resolve both regions. The optimized stations are spaced very far from current SGS stations; such a station distribution is beneficial to accurately locate earthquake hypocenters, and thus will help to study seismicity patterns and to better understand the underlying seismotectonics.

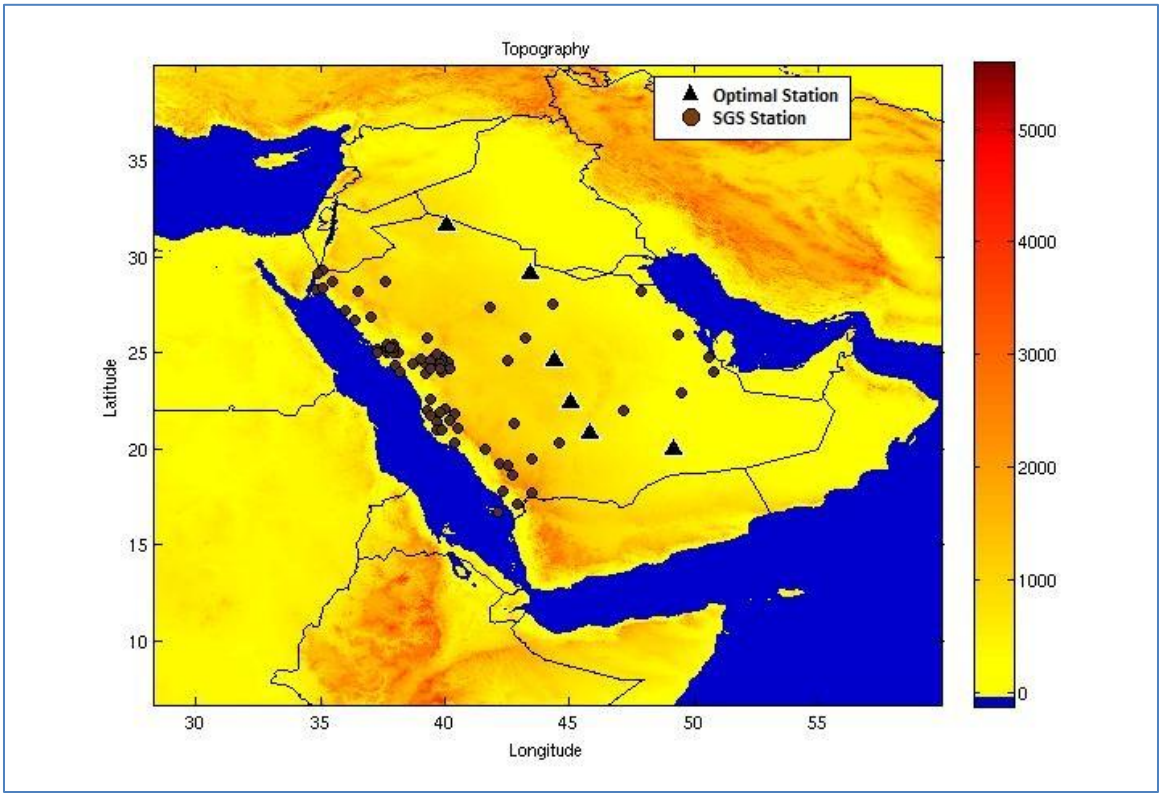

Figure 19: The updated Saudi network. Six station locations (black triangles) were optimized and plotted along with the existing SGS stations (brown circles). The color shading depicts topography.

Table 4: Proposed Stations Coordinates

Station Number	Latitude [^{o}]	Longitude [^{o}]
1	31.6646	40.0677
2	29.1652	43.4013
3	24.5830	44.3833
4	22.4502	45.0381
5	20.8340	45.8229
6	20.0008	49.1653

To understand better the optimized station locations, we display the earthquake probability distribution with the optimized and the existing SGS stations (Figure 20). The curvature of the optimized stations in the north and in the south follows the trend outlined by the regions of high earthquake probability.

The velocity model and the earthquake probability distribution play an essential role in optimizing the network. Thus, if one wanted to increase the significance of a certain seismically active region only the probability in that zone needs to be increased. Naturally, as the network expands and is updated we will know more about the seismicity and the geology. This will help us develop more accurate models to further update the network.

After obtaining the initial results for the optimized station distribution, the optimization for six stations was repeated for a different initial station configuration. We notice that

the new optimized network has 5 of the stations almost identical in position while one of the stations in the south (station 5) tends to move 100 km to the south.

Our technique is very sensitive to the earthquake probability distribution. Before we evaluated the final probability distribution we tested this method on different earthquake probability distributions, the station configuration varied widely.

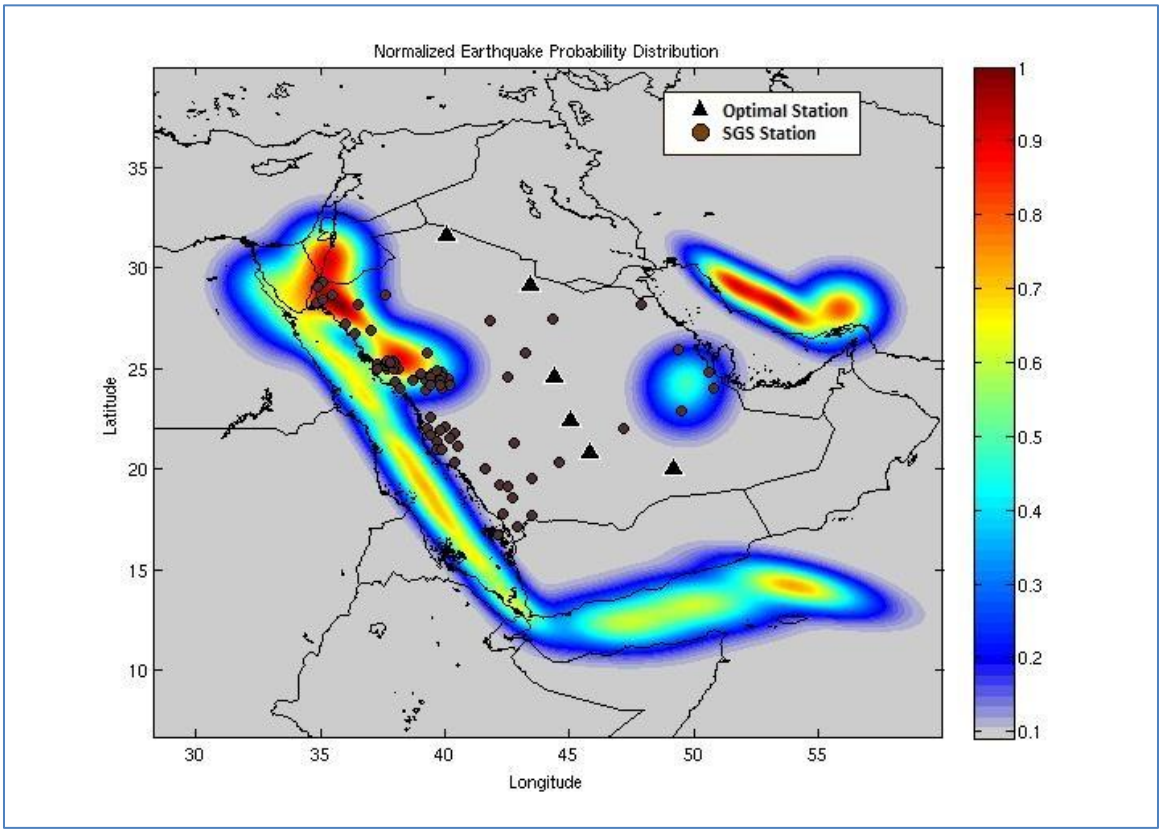

Figure 20: The same optimized station as shown in Figure 18 (black triangles) and the SGS stations (brown circles) are plotted against the earthquake probability distribution.

4.2 Analysis of the Network

The planned update of the SGS network serves to reduce the errors in estimating hypocenter locations. To illustrate this argument, we calculated and map the standard errors of the epicenter position (σ_{xy}) and hypocenter depth (σ_{zo}). Following the derivation of Kijko (1977-b), the model covariance matrix for the hypocenter parameters may be described as:

$$D[< \theta >] = (A^T W A)^{-1}, \tag{4.1}$$

where:

$$A_{ij} = \frac{\partial f(X_i, \theta)}{\partial \theta_j} \tag{4.2}$$

$$W_{ij} = \begin{Bmatrix} \frac{1}{\sigma_t^2}, i = j \\ 0, i \neq j \end{Bmatrix} \tag{4.3}$$

Thus

$$d[<\theta>]_{ii} = \left\{(A^T W A)_{ii}^{-1}\right\} = \begin{Bmatrix} \sigma_{t_o}^2(\theta'), i = 1 \\ \sigma_{x_o}^2(\theta'), i = 2 \\ \sigma_{y_o}^2(\theta'), i = 3 \\ \sigma_{z_o}^2(\theta'), i = 4 \end{Bmatrix} \quad (4.4)$$

Where W is residual pick time standard error at the ith station. $\sigma_{t_o}, \sigma_{x_o}, \sigma_{y_o}, \sigma_{z_o}$ are the calculated standard errors associated with each hypocenter parameter. This allows to us map the standard errors of all relevant parameters. To understand the standard error associated with epicenter position we define the following term (Kijko 1977-b).

$$\sigma_{xy}(\theta') = \{d[<\theta>]_{22} d[<\theta>]_{33} - (d[<\theta>]_{23})^2\}^{1/4} \quad (4.5)$$

Using Equation 4.1-4.5 we investigate the errors in determining the hypocenter depth and the epicenter location for a certain variance of the time residual for the *i*th station (Kijko and Sellevoll 1982). Ideally, each station should have a certain variance in time residuals. In principle, travel time residuals may change from station to station, and are (most likely) correlated between neighboring stations.

However, because we did not have any waveform data, we estimated the standard error of the residual time to be constant for stations with certain distance from the epicenter. We tested different values for the residual time standard error (0.1 and 0.05 seconds) in order to determine which travel-time residual provided realistic epicentral

position and hypocenter depth through their respective standard errors. The value of the standard error of time residual that we chose to use is 0.075 seconds for stations less than 100 km and 0.15 seconds for stations more than 100 km from the epicenter. This is because the near stations will have a greater signal-to-noise ratio than the farther stations, thus they will generally have lower standard error than the farther stations. We chose 100 km because according to Richter definition of magnitude it can be assumed that with standard seismograph magnitude events 3 can be recorded at 1 micron at 100 km, which should be sufficiently above the noise level

Figure 21 shows the standard error of earthquake depth before adding the optimized stations. The standard error of earthquake depth is low (about 1-4 km) near the western border of the kingdom. But in the north, east, and south regions the error rises significantly due to the lack of stations. When analyzing the standard errors we limit the maximum error to 25 km, to better resolve the small standard variations for earthquakes originating inside Saudi Arabia

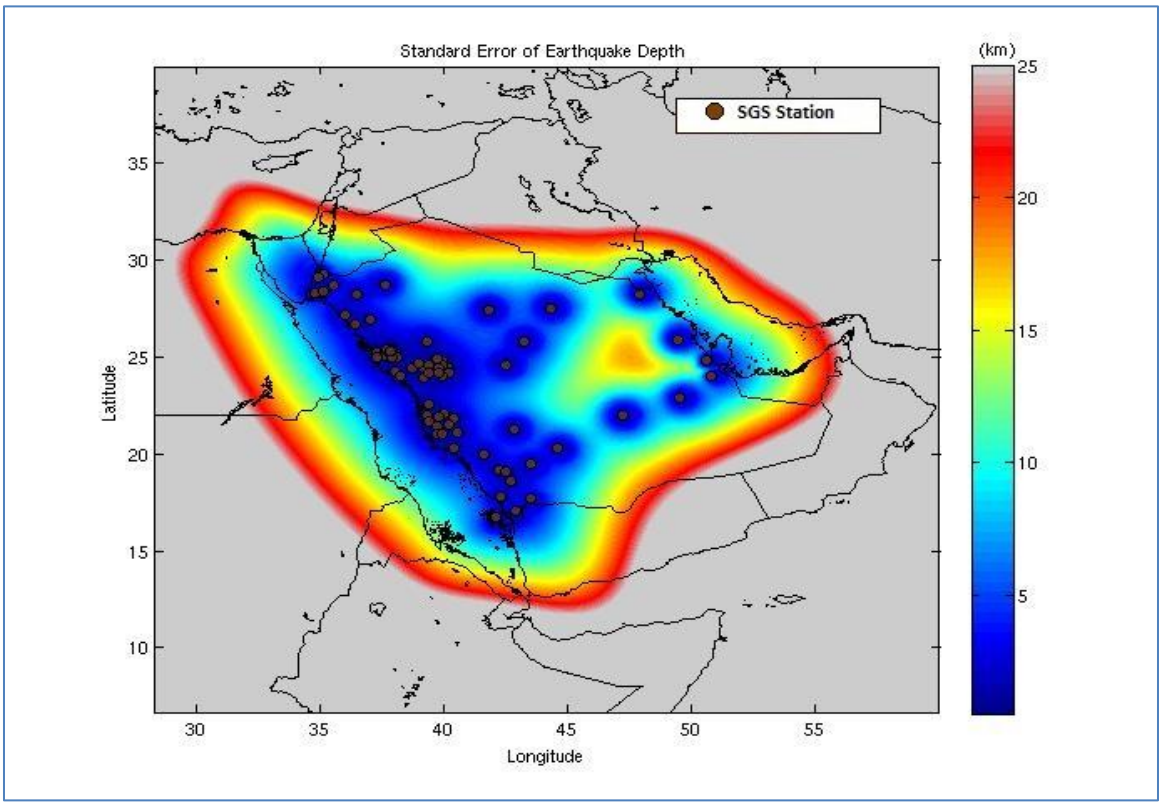

Figure 21: The standard error of the earthquake depth for the current network, notice how the depth is poorly constrained in the northern and southern parts.

Additionally we analyzed the standard error of the epicenter position (Figure 22).

Epicentral uncertainty in Saudi Arabia is smaller than the hypocenter depth uncertainty.

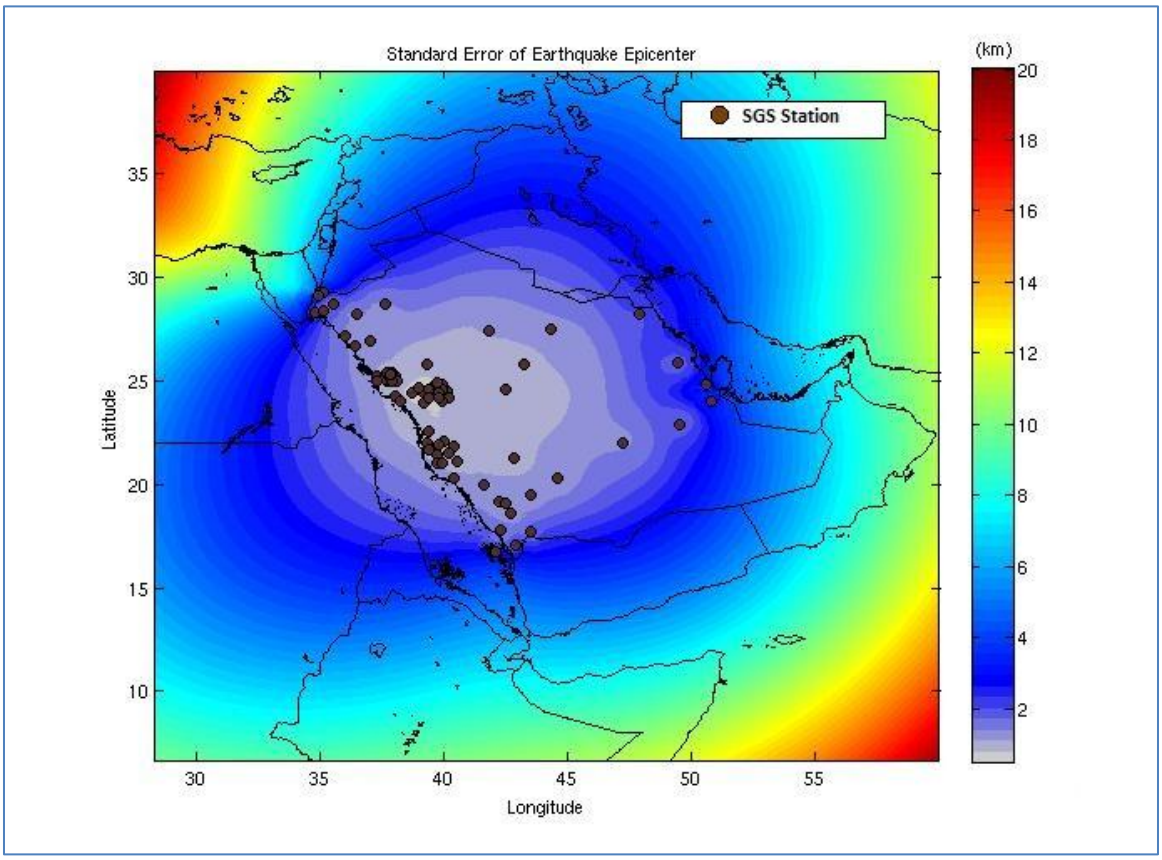

Figure 22: The standard error of epicenter location for the current network.

After adding the optimized stations the standard errors were recomputed. Figure 23 and 24 show the standard error of earthquake depth and for epicenter location, respectively, after adding the optimized stations. Notice that the depth is much better constrained in the north, south, and east. As result the low error contours extend farther in the Kingdom. Only the southern eastern region of the Kingdom remains of a high standard error for earthquake depth; however this region is not subjected to high seismicity compared to the West Coast of the Kingdom. Adding the proposed station the area of the low error contours tends to get larger.

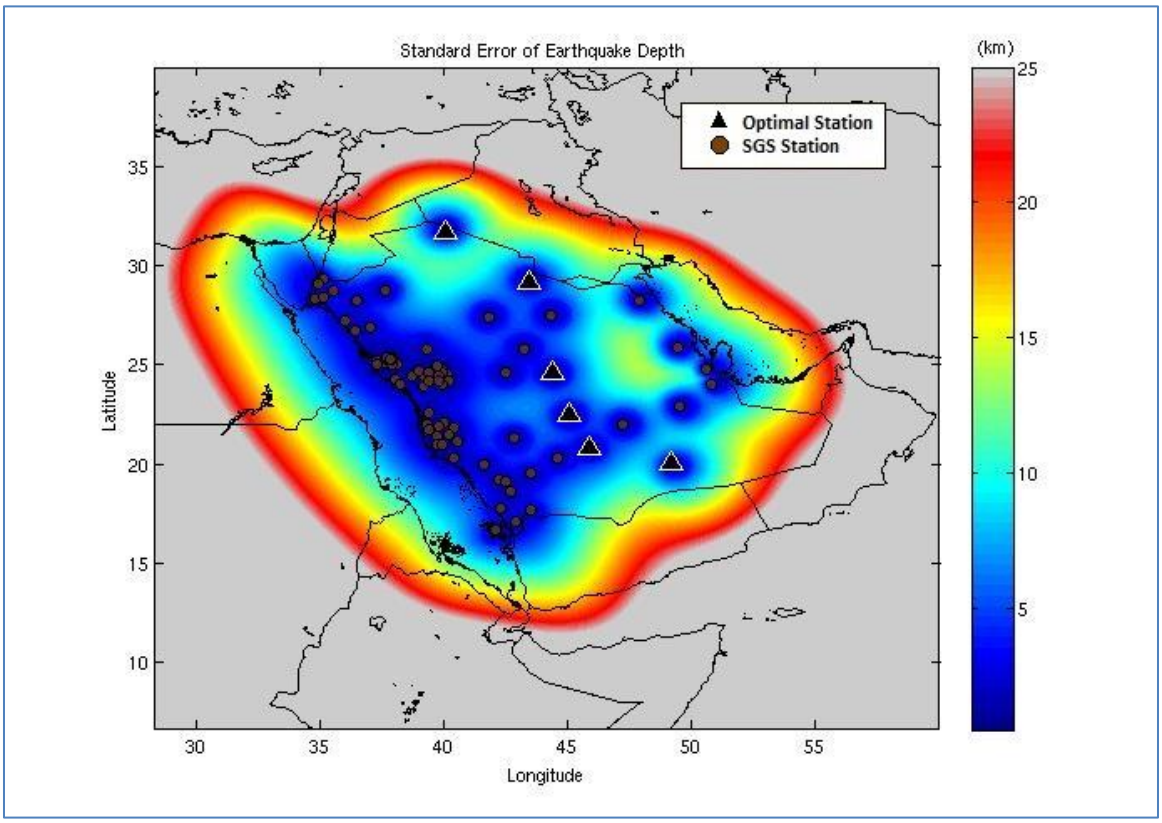

Figure 23: The standard error of earthquake depth with the SGS (brown circles) and the proposed stations (black triangles) overlain.

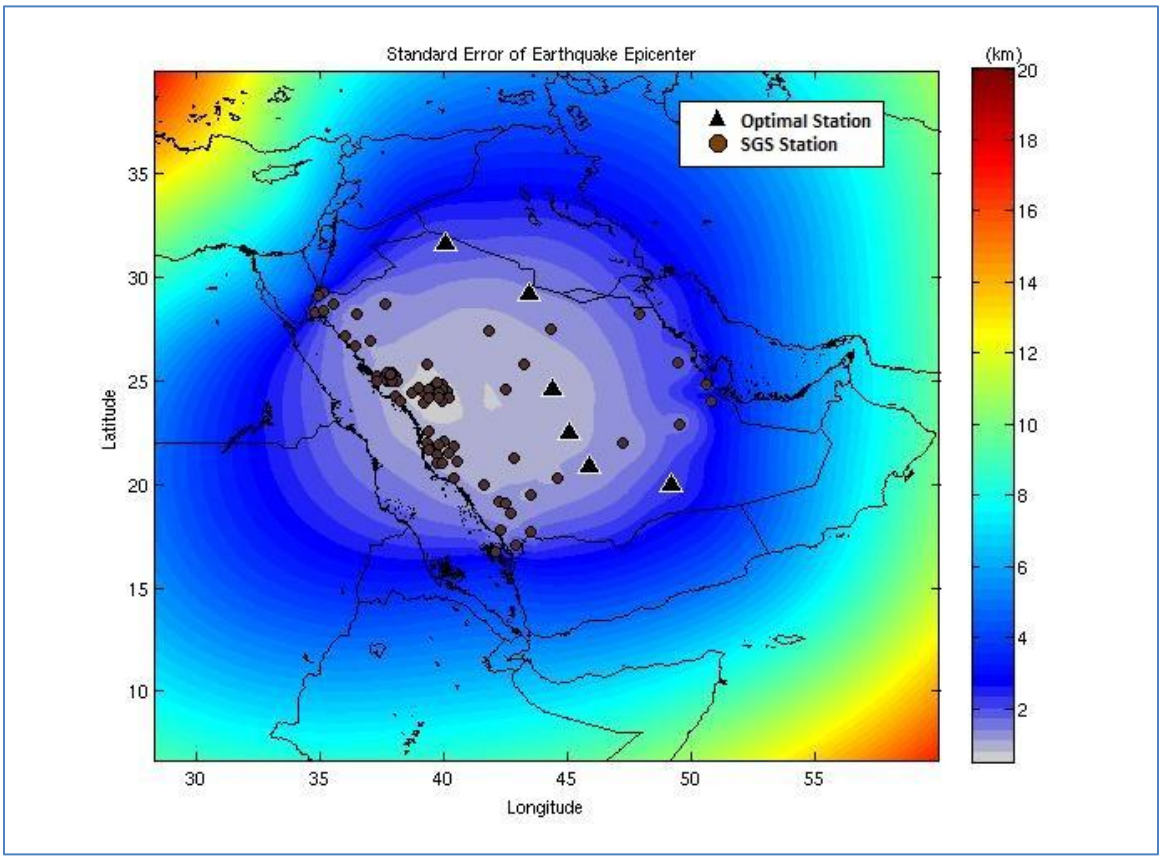

Figure 24: The standard error of epicenter position after adding the purposed stations to the current SGS stations.

After obtaining the optimum station locations for six stations, we optimized station locations again with a different number of stations to better understand the effect of increasing number of stations on the standard error of hypocenter parameters. We then calculated the mean error of hypocenter parameters (x_o,y_o,z_o) inside Saudi Arabia and plotted the smoothed mean error versus the number of stations to optimize for (Figure 25). We notice that there is a trend, the mean errors decreases with the increase of number of stations.

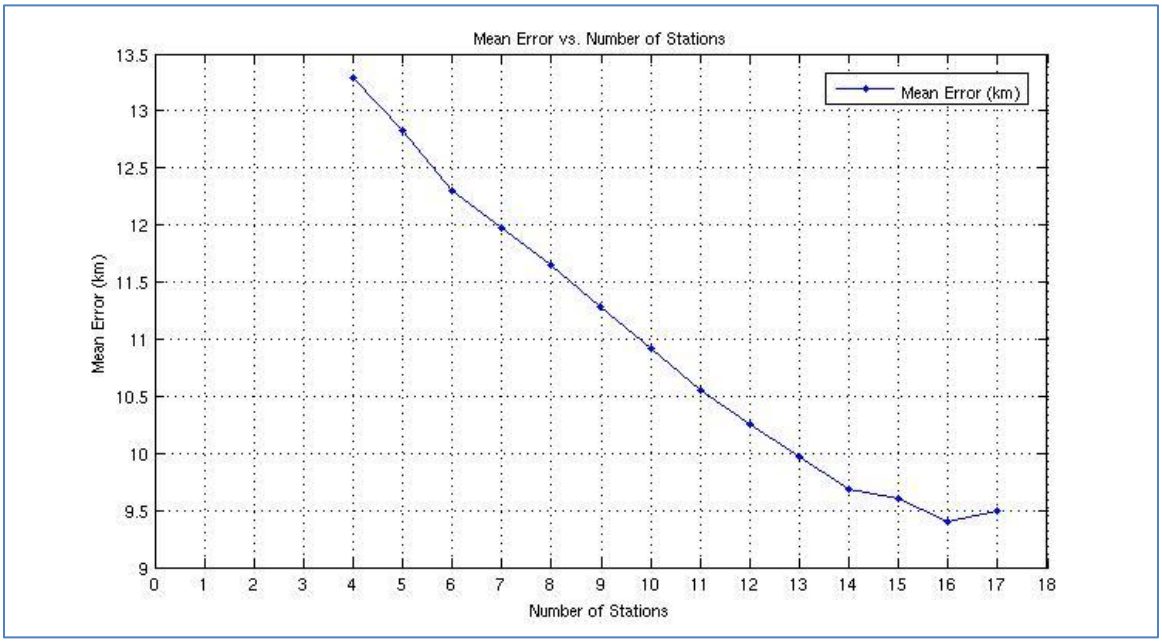

Figure 25: The smoothed mean error of hypocenter parameters in Saudi Arabia versus the number of stations used to optimize.

4.3 Future Work

Currently we are working with the SGS to improve their hypocenter location capabilities within a smaller region in Saudi Arabia. This project is being carried out, in association with the SGS, in two regions of high seismicity inside the Kingdom that pose a threat to the surrounding infrastructure. The resulting optimized station locations will be more accurate since the velocity model is more accurate. Also, geologically mapped faults in the area will help in defining the earthquake probability distribution more precisely.

Another interesting issue we are analyzing is the efficiency of the network versus the number of stations used. It is known that as the number of stations increases the efficiency tends to saturate (Kijko 1977-b). By examining further we will be able to suggest optimum location and number of stations to update the current network.

Finally, we are developing a similar method that uses the basic theory presented here and applies it for designing the optimum network around an injection well to monitor the micro-seismic events. By locating their hypocenter we will be able to understand better the fluid flow in the subsurface and monitor reservoir parameters.

Chapter 5

Conclusion

Our optimization method and its improvements provide a useful tool for designing the optimal locations of new stations in the Saudi network. The deployment of these proposed station sites will enable seismologists to locate hypocenter coordinates with increased accuracy, which will help in better understanding the regional seismicity and the neo-tectonics with respect to the geological boundaries. Our improvements, using the “D-optimum Criterion”, in particular account for direct and refracted waves and their decreasing signal-to-noise ratio (due to geometric spreading) for far-distant stations. The additional benefit from this approach is that it facilitates tomographic studies using secondary phases to better image Earth structure.

The method is very flexible which helps in designing new networks or updating existing networks. Another advantage is that the method is not affected by the size of the network, and therefore can be applied for large scale and small scale regions.

In our case we updated an existing network (SGS network) with six broadband stations. The station locations where placed in a curved line that runs from north to south in the central part of the Kingdom. The line curves around (concave towards) regions of high seismicity.

Analyzing the standard errors of hypocenter depth shows that the errors were decreased in the northern and southern areas. It is well constrained for the western part

of the Kingdom due to the large number of stations deployed there. On the other hand the epicenter position errors were quite low in the entire Arabian peninsula but increases dramatically when moving away from the peninsula, due to the limited azimuthal coverage obtained from the events at Saudi borders.

REFERENCES

Al-Amri, A. M., A. J. Rodgers, et al. (2008). "Improving the level of seismic hazard parameters in Saudi Arabia using earthquake location." Arabian Journal of Geosciences **1**(1): 1-15.

Al-Amri, A. M. (1995). "Recent Seismic Activity in the Northern Red-Sea." Journal of Geodynamics **20**(3): 243-253.

Al-Amri, A. M.,1998, "The crustal structure of the western Arabian Platform from the spectral analysis of long-period P-wave amplitude ratios." Tectonophysics **290**, 271-283.

Benoit, M. H., A. A. Nyblade, et al. (2003). "Upper mantle P wave velocity structure and transition zone thickness beneath the Arabian Shield." Geophysical Research Letters **30**(10).

Box, G. E. P. and H. L. Lucas (1959). "Design of Experiments in Non-Linear Situations." Biometrika **46**(1-2): 77-90.

Camp, V. E. and M. J. Roobol (1992). "Upwelling Asthenospear Beneath Western Arabia and its Regional Implication." Journal of Geophysical Research-Solid Earth **97**(B11): 15255-15271.

Cochran, J. R. (1981). "The Gulf of Aden - Structure and Evolution of a Young Ocean-Basin and Continintal-Margin." Journal of Geophysical Research **86**(NB1): 263-287.

Debayle, E., J. J. Leveque, et al. (2001). "Seismic evidence for a deeply rooted low-velocity anomaly in the upper mantle beneath the northeastern Afro/Arabian continent." Earth and Planetary Science Letters **193**(3-4): 423-436.

Engdahl, E. R., J. A. Jackson, et al. (2006). "Relocation and assessment of seismicity in the Iran region." Geophysical Journal International **167**(2): 761-778.

Garcia-Fernandez, M., A. Kijko, et al. (1988). "Optimum Station Distribution to Monitor Seismic Activity of Teide Volcano, Tenerife, Canary Islands." Journal of Volcanology and Geothermal Research **35**(3): 195-204.

Ghalib, H. A. A., D. R. Russell, et al. (1984). "Optimal-Design of a Regional Seismological Network for The Arab Countries." Pure and Applied Geophysics **122**(5): 694-712.

Hardt, M. and F. Scherbaum (1994). "The Design of Optimum Networks for Aftershock Recordings." Geophysical Journal International **117**(3): 716-726.

Howells, D. A. (1983). "A History Of Persian Earthquakes - Ambraseys,Nn, Melville,Cp." Earthquake Engineering & Structural Dynamics **11**(4): 591-591.

Jackson, J. and T. Fitch (1981). "Basement Faulting and the Focal Depths Of The Larger Earthquakes in the Zagros Mountains (Iran)." Geophysical Journal of the Royal Astronomical Society **64**(3): 561-586.

Kijko, A. (1977-a). "Algorithm for Optimum Distribution of a Regional Seismic Network .1." Pure and Applied Geophysics **115**(4): 999-1009.

Kijko, A. (1977-b). "Algorithm for Optimum Distribution of a Regional Seismic Network .2. Analysis of Accuracy of Location of Local Earthquakes Depending on Number of Seismic Stations." Pure and Applied Geophysics **115**(4): 1011-1021.

Kijko, A. and M. A. Sellevoll (1982). "A Statistical-Model for Estimating The Accuracy of Event Location Applied to a Network Of Scandinavian Stations." Pure and Applied Geophysics **120**(1): 186-196. 56

Klinger Y, Rivera L, Haessler H, Maurin J-C (1999). “Active faulting in the Gulf of Aqabah: New knowledge from the MW 7.3 earthquake of 22 November 1995.” Bulletin of the Seismological Society of America **89**:1025–1036

Lee, W. H. K. and S. W. Stewart (1981). Principles and applications of microearthquake networks, Academic Press.

Lilwall, R. C. and A. Douglas (1970). "Estimation of /b P/-wave travel times using the Joint Epicentre Method." Geophysical Journal of the Royal Astronomical Society|Geophysical Journal of the Royal Astronomical Society **19**(2): 165-181.

Menke, W. (1984). "Geophysical data analysis. Discrete inverse theory." Geophysical data analysis. Discrete inverse theory: xii+260.

Oth, A., M. Bose, et al. (2010). "Evaluation and optimization of seismic networks and algorithms for earthquake early warning - the case of Istanbul (Turkey)." Journal of Geophysical Research-Solid Earth **115**.

Park, Y., A. A. Nyblade, et al. (2007). "Upper mantle structure beneath the Arabian Peninsula and northern Red Sea from teleseismic body wave tomography: Implications for the origin of Cenozoic uplift and volcanism in the Arabian Shield." Geochemistry Geophysics Geosystems **8**.

Rabinowitz, N. and D. M. Steinberg (1990). "Optimal configuration of a seismographic network - a statistical approach." Bulletin of the Seismological Society of America **80**(1): 187-196.

Rodgers, A. J., W. R. Walter, et al. (1999). "Lithospheric structure of the Arabian Shield and Platform from complete regional waveform modelling and surface wave group velocities." Geophysical Journal International **138**(3): 871-878.

Schorlemmer, D. and J. Woessner (2008). "Probability of detecting an earthquake." Bulletin of the Seismological Society of America **98**(5): 2103-2117.

Stein, S. and M. Wysession (2003). An introduction to seismology, earthquakes, and earth structure. Malden, MA, Blackwell Pub.

Stocklin, J. (1968).”Structural history and tectonics of Iran: a review.” Bulletin of the Seismological Society of America **52**: 1229-1258.

Torczon, V. (1997). "On the convergence of pattern search algorithms." Society for Industrial and Applied Mathematics Journal on Optimization **7**(1): 1-25.

Uhrhammer, R. A. (1982). "The optimal estimation of earthquake parameters." Physics of the Earth and Planetary Interiors **30**(2-3): 105-118.